MAY 25 2004

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	24.May.04	DISSERTATION

4. TITLE AND SUBTITLE
LUMINOUS EFFICIENCY FUNCTIONS AT HIGHER INTENSITIES

5. FUNDING NUMBERS

6. AUTHOR(S)
LT COL HARRINGTON LAWRENCE K

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
UNIVERSITY OF MISSOURI ST LOUIS

8. PERFORMING ORGANIZATION REPORT NUMBER
CI04-352

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
THE DEPARTMENT OF THE AIR FORCE
AFIT/CIA, BLDG 125
2950 P STREET
WPAFB OH 45433

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
Unlimited distribution
In Accordance With AFI 35-205/AFIT Sup 1

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

20040601 032

14. SUBJECT TERMS	15. NUMBER OF PAGES
	63
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

THE VIEWS EXPRESSED IN THIS ARTICLE ARE THOSE OF THE AUTHOR AND DO NOT REFLECT THE OFFICIAL POLICY OR POSITION OF THE UNITED STATES AIR FORCE, DEPARTMENT OF DEFENSE, OR THE U.S. GOVERNMENT.

Luminous Efficiency Functions at Higher Intensities

Lawrence Kent Harrington

Dissertation Submitted for the Degree Doctor of Philosophy in Physiological Optics
University of Missouri-St. Louis

Abstract

Two psychophysical measurement techniques, flicker photometry and successive heterochromatic brightness matching, were used to measure changes in luminance efficiency functions with increasing levels of light adaptation. Both measurement techniques were performed using the same optical system and the same seven healthy adults as subjects. Measurements were taken at four reference stimulus intensities, 1, 10, 100 and 1000 foot-lamberts. Luminous efficiency was found to depend on both the technique and the reference stimulus intensity with which the measurements were taken. For heterochromatic brightness matching, luminous efficiency increased for longer wavelengths as reference intensity increased. Peak luminous efficiency shifted from approximately 540nm to greater than 600nm with increasing intensity for all seven subjects. Peak luminous efficiency was constant for flicker photometry across all intensities but the function narrowed slightly at 100 foot-lamberts.

Luminous Efficiency Functions at Higher Intensities

Dissertation
Submitted for the Degree
Doctor of Philosophy in
Physiological Optics

University of Missouri-St. Louis

Lawrence Kent Harrington
April 2004

Advisory Committee

Carol Peck, PhD
Carl Bassi, PhD
W. Gary Bachman, OD, MS
Frank Moss, PhD

Table of Contents

Abstract

Two psychophysical measurement techniques, flicker photometry and successive heterochromatic brightness matching, were used to measure changes in luminance efficiency functions with increasing levels of light adaptation. Both measurement techniques were performed using the same optical system and the same seven healthy adults as subjects. Measurements were taken at four reference stimulus intensities, 1, 10, 100 and 1000 foot-lamberts. Luminous efficiency was found to depend on both the technique and the reference stimulus intensity with which the measurements were taken. For heterochromatic brightness matching, luminous efficiency increased for longer wavelengths as reference intensity increased. Peak luminous efficiency shifted from approximately 540nm to greater than 600nm with increasing intensity for all seven subjects. Peak luminous efficiency was constant for flicker photometry across all intensities but the function narrowed slightly at 100 foot-lamberts.

Introduction

Radiometry is the measurement of radiant, electromagnetic, energy. For optical radiation it encompasses techniques for counting the number of photons per unit time, area, or space. The challenge in radiometry is that the size, velocity and number of photons make it impossible to count photons directly. Therefore we must rely on indirect measures such as observing and quantifying the optical radiation effects on an absorptive material. Specifically, radiometry depends on the absorption of radiant energy producing a measurable electromagnetic, chemical, or thermal response.

This indirect measurement of radiation would be straight forward if it were not for one major complicating factor. The photon's ability to generate these measured responses is dependent on wavelength. The transmission of optical media between the light source and the detector, the absorption of the detector and the amount of energy converted to the measured response all vary with wavelength. To compensate for wavelength dependence, a correction factor is typically calculated for each nanometer of wavelength. The aggregation of these correction factors across the spectrum of concern constitutes a spectral weighting function.

Photometry is the measurement of light, the radiant electromagnetic energy that is useful to human vision. Interest in the accurate measurement of light grew dramatically with advances in the fields of astronomy and artificial lighting in the last part of the 19th century (Johnston, 2001). Early photometry techniques involved the visual comparison of the test light against a known light standard. The photometrist would match the brightness, produced on a surface, by the test light to the brightness produced on the same surface by a standard light, such as a candle or an oil lamp. The brightness match was

made by adjusting the distance of the test light from the surface. This method of adjusting the relative distance of the test light to the standard light led directly to several older units of illuminance such as the foot-candle, however, it had several limitations. The standard lights were inconsistent and difficult to maintain. Their luminance varied with environmental conditions and special attention had to be given to such factors as candle wax purity, wick length, ambient temperature, humidity and atmospheric pressure. Individual observer variation also contributed a substantial amount of error to light measurements. The human error in the measurement was particularly large when the test light and standard were of different colors. To reduce cost and error, a photometric technique was needed that eliminated both the human observer and the standard light from the measurement process. If a spectral weighting function could be developed for the human eye, the measurement of light could be reduced to a simple mathematical transformation of a radiometric measurement. This desired spectral weighting function for vision is often referred to as the luminous efficiency function.

In 1924 the Commission Internationale de l'Eclairage (CIE), in acknowledgment of a very practical need for a spectral weighting function for human vision, adopted the CIE spectral luminous efficiency function for photopic vision $V(\lambda)$. The luminous efficiency function $V(\lambda)$, based on an amalgam of flicker photometry and step-by-step brightness matching data collected from over two hundred and fifty subjects, has been used to great advantage by vision scientists and lighting engineers and has revolutionized the technical field of photometry (Gibson and Tyndall, 1923; Wyszecki and Stiles, 1982).

The development of electrically calibrated radiometers and the availability of $V(\lambda)$ dramatically improved the reliability and applicability of photometric techniques

(Wyszecki and Stiles, 1982). Now using a radiometer to measure radiant power (P) in watts, luminous flux (F) in lumens can be calculated using the integral:

$$F = 683 \int V(\lambda)P(\lambda)\, d\lambda$$

More practically, total luminous flux can be approximated by the summation of the calculated luminous flux at each wavelength.

$$F = \Sigma\, 683\, V(\lambda)P(\lambda)$$

Intrinsic in these equations is the idea that the brightness produced by a given amount of luminous flux is fixed and independent of the spectral content. Different authors describe this assumption in various ways, often by briefly mentioning additivity or Abney's Law. Wyszecki and Stiles (1982) give a comprehensive discussion of the "basic laws of brightness matching" including the laws of symmetry, transitivity, proportionality and additivity. However, exceptions to the laws have been discovered, explored and used to help explain visual processes.

The classic exception concerns visual performance under low levels of illumination (scotopic conditions). $V(\lambda)$ and the brightness matching laws do a poor job of predicting visual performance under low light conditions. Investigation of this failure led to a better understanding of neural physiology and visual processing. The failure of $V(\lambda)$ and the laws of brightness matching under scotopic conditions was largely explained by the sensitivity range of retinal photoreceptors and ultimately a second spectral luminous efficiency function for scotopic vision ($V'\lambda$) was established in 1951.

There are other conditions where the visual performance departs from $V(\lambda)$ in ways that are more difficult to explain. The relative sensitivity of the eye also varies with target size, retinal eccentricity and target duration (Ikeda et al, 1982; King-Smith and Carden, 1976). Not all of these irregularities can be explained at the photoreceptor level. The explanations for non-photoreceptor caused departures of relative sensitivity from $V(\lambda)$ have generally settled on the idea that the perception of brightness is determined by two neural systems, a chromatic system and an achromatic system. The relative sensitivity of the eye at any particular time is dependent on the relative contributions of these two systems. As Meyer (1978) noted

> This explanation states that the output from cones feeds into two systems, one spectrally opponent or chromatic and the other one achromatic. Signals from the cones to the non-opponent achromatic system are combined linearly and activity at higher neural levels can be accurately predicted from the sum of the inputs. Signals for the cones to the opponent or chromatic system however are antagonistic in that activity within the red-green system (or the blue-yellow) is subtracted from one another. Thus, the specific luminous efficiency function obtained in a given experiment depends upon whether the method used taps the output of the achromatic system (flicker, minimum border, etc.) or the outputs of both the chromatic and achromatic systems (heterochromatic brightness matching, absolute foveal threshold, etc.)

Therefore we are left with the difficult situation where the relative sensitivity of the eye or the luminous efficiency of a light source is not fixed but fluctuates with the viewing condition. There are two approaches to dealing with this situation. Photometry can be treated as a physical measurement that may or may not correlate well with the perception of brightness or $V(\lambda)$ can become one part of a system of luminous efficiency curves used in photometry with the viewing condition dictating which curve is most appropriate for the given application.

In response to this problem the CIE acknowledged that the meaningful quantification of light often requires more than the indiscriminant use of a V(λ)-corrected physical photometer. An understanding of the visual system is required to select the most appropriate spectral luminous efficiency function for a given situation.

> For photopic vision and luminances larger than several cd/m^2, ordinary physical photometers corrected to V(λ) give visually accurate measures for small, centrally fixated lights of broad spectral composition. For all other applications a different luminous efficiency function should be employed (Meyer et al, 1978).

However, knowledge of the limitations of photometric techniques does not necessarily result in improved measurements. Photometry has a recurring history of imprecise results leading to a poor reputation and low expectations, where the practitioners either "accepted what they recognized as an imprecise measurement or carried on unaware of the potential systematic errors." (Johnston, 2001) This rather negative assessment of the practice of photometry may describe our current ability to measure some of the newer light sources such as LEDs and lasers.

LEDs and lasers are spectrally narrow sources and as such are known to be problematic for the V(λ) based photometric system (Meyer et al, 1978; Kinney, 1983). The search for an alternative luminous efficiency function leads us back to the psychophysical measures for deriving human spectral weighting functions.

There are several psychophysical measures that can be used for deriving a system of relative sensitivity functions including simultaneous brightness matching, successive brightness matching, the minimally distinct border method, grating visual acuity and increment threshold. These measures have been described thoroughly elsewhere;

successive heterochromatic brightness matching has been described by Ikeda and Shimozono (1978) and the other techniques by Wyszecki and Stiles (1982). The resulting relative sensitivity functions derived from these methods fall generally into two categories.

The first category includes spectral luminous efficiency functions generated by flicker photometry, grating visual acuity and minimally distinct border. These functions are thought to be representative of the achromatic system, which is derived from the linear sum of outputs from the middle and long wavelength cones. They are traditionally thought to be independent of any neural contributions from the short wavelength cone and chromatic pathways because of inferior temporal (flicker photometry) or spatial (visual acuity, minimally distinct border) characteristics. (Lennie et al, 1993) They tend to be relatively narrow and smooth with peak sensitivity at approximately 555 nm. Most importantly these luminous efficiency functions are congruous with the brightness matching laws required to support a meaningful photometric system.

The second category of functions includes heterochromatic brightness matching and increment threshold spectral luminous efficiency functions. These functions are thought to be mediated by both the chromatic and achromatic systems. They have a broader, flatter curve, sometimes demonstrating a multiple peaked appearance (Ikeda and Yaguchi, 1982; Sperling and Harwerth, 1971).

Intuitively, brightness matching seems the logical psychophysical function for generating a luminous efficiency function (Wagner and Boynton, 1972; Meyer et al, 1978) because the experimental conditions are representative of normal visual experience. Unfortunately brightness matching suffers from reliability problems and the

results obtained with heterochromatic brightness matching fail to observe the brightness matching laws. This was one of the major reasons that flicker photometry, not heterochromatic brightness matching, played such a major role in the development of $V(\lambda)$. However as mentioned earlier, $V(\lambda)$ and flicker photometry do a poor job of predicting brightness perception for spectrally narrow sources. A luminance efficiency function based on heterochromatic brightness matching may be more appropriate, then for photometric measurement of LEDs and lasers. However, these light sources can produce retinal illuminances several orders of magnitude higher than those used to generate most of the published heterochromatic brightness matching luminance efficiency curves (Bedford and Wyszecki, 1958; Comerford and Kaiser, 1975; Guth and Lodge, 1973; Kinney 1964; Sperling and Lewis, 1959; Wagner and Boynton, 1972; Sliney and Wolbarsht, 1980).

Whether retinal illuminance is an important variable in determining the appropriate luminous efficiency function for a given set of photopic conditions is controversial. The CIE provides extensive guidance on the impact of low (scotopic and mesopic) luminance levels on luminous efficiency functions in Light as a True Visual Quantity: Principles of Measurement but fails to mention any impact of luminance levels within the photopic range. Deane Judd, well known for his work leading to the establishment of $V_M(\lambda)$ (CIE 1988 Modified 2 Degree Spectral Luminous Efficiency Function for Photopic Vision) (Kaiser 1990), concluded that $V(\lambda)$ extrapolates well to very bright sources. In 1951 Judd stated that $V(\lambda)$ was applicable for the luminance range starting around 1 foot-lambert and going up to 10,000 foot-lamberts (Stevens, 1951). However, changes in luminous efficiency across photopic luminance levels have

been documented for many years. In 1909 Helmholtz, in his Treatise on Physiological Optics, was one of the first authors to write about the dependence of relative sensitivity on retinal illuminance, saying:

> When whites of different luminosity are obtained by mixture, the amounts of the complementary colours are in a constant ratio to each other in objective intensity, but in a very variable ratio to each other in subjective luminosity (Southall, 1962).

Increment threshold experiments provide some of the most convincing evidence for changes in luminous efficiency functions over photopic luminance levels. Sperling and Harwerth found a dramatic change in the shape of threshold luminous efficiency curves depending on background luminance. As they increased the background luminance, from 0 to 10,000 trolands, the luminous efficiency curves developed three peaks, a broad peak around 430 nm and two more narrow peaks around 535nm and 610nm. They explained these three peaks as an increase in the interaction between the long and middle wavelength cones associated with an increase in the adaptive state of the fovea (Sperling and Harwerth, 1971). Since a chromatic channel explanation was given for this luminous efficiency change, it is logical to look for a similar change in other luminous efficiency measures that tap into the chromatic channel, such as heterochromatic brightness matching.

As was the case with increment threshold derived luminous efficiency curves, there is evidence that heterochromatic brightness matching luminous efficiency curves vary with retinal illuminance. Some authors have found a relative loss of visual sensitivity to red stimuli with increasing luminance levels in the low photopic region, a

phenomenon that has been referred to as a reversed or inverse Purkinje shift (Thomson, 1946; Yaguchi and Ikeda, 1980; Bedford and Wyszecki, 1958). Others found that not only were the heterochromatic brightness matching curves broader than V(λ) but for some of their subjects, the difference between the two functions increased with increasing retinal illuminance (Yaguchi and Ikeda, 1980; Sagawa et al, 1991). Wagner and Boynton (1972) conducted a series of heterochromatic photometry experiments and found that their data deviated substantially from their predictions. They explained the deviation by a failure to hold luminance constant at long and short visible wavelengths due to the limited radiance produced by their light source.

Wyszecki and Stiles used trichromatic color matching, a close relative of brightness matching, to model pigment bleaching at higher retinal illuminances. They found that the relative contributions of long, middle and short wavelength lights required to match a white reference varied dramatically with reference retinal illuminance between 8,000 and 50,000 trolands (Wyszecki and Stiles, 1980). This breakdown of the law proportionality is important to both color matching and brightness matching and the favored physiological interpretation for the breakdown is the asymmetric bleaching of cone photopigments (Wyszecki and Stiles, 1982).

The case for changes in the flicker photometry luminous efficiency curves with increasing luminance is less convincing, although several authors have found small changes (De Vries, 1948; Sagawa et al, 1991). Most notably, the researchers from whose work V(λ) was derived felt that luminance level contributed substantially to the shape of the luminous efficiency curve found by flicker photometry (Ives, 1912; Gibson and Tyndall, 1923).

Having reviewed the history of photometry it seems naïve to assume that a photometric system based on $V(\lambda)$ is going to produce reliable predictions of visual performance in the case of bright and spectrally narrow light sources like LEDs and lasers. Therefore I decided to investigate luminous efficiency functions at high intensities using two psychophysical techniques: One that is representative of the visual performance based on achromatic system alone, namely flicker photometry and one that would tap into both the achromatic and chromatic systems, successive heterochromatic brightness matching. Successive heterochromatic brightness matching is felt to produce similar results to the traditional bipartite heterochromatic brightness matching and has the advantage of using the same apparatus as flicker photometry (Ikeda and Shimozono, 1978). The major goal was to compare results of the 2 methods directly, using the same optical system and subjects in order to explore how luminance efficiency functions change with increasing intensity so that we can better predict the visual perception of spectrally narrow and bright light sources.

Methods

Apparatus

Light from a 1000-watt xenon arc lamp was split 99/1 using an antireflective window to produce two illumination channels, a spectrally broad (white) reference channel and a monochromatic test channel. The test channel, which consumed the majority of the lamp output, could be varied in intensity and monochromatic spectral content using an iris aperture, a remotely controlled, motor driven, duel, counter rotating set of variable density filters (variable density filter system) and a motorized filter wheel

which contained 17 narrow (10mm full width-half maximum) bandpass interference filters. The reference channel was modulated in intensity using a circular variable density filter and an iris aperture. A rotating mirrored, variable frequency optical chopper was used to merge the two channels spatially while maintaining temporal separation. The reference and test channels alternately illuminated the end of a .75 in diameter acrylic cylinder that served both as a diffusing optic to dampening the variation in beam intensity and a back lit viewing screen to produce the stimulus field. A chin rest was used to position the subject 21 inches from the viewing end of the diffusing optic. The stimulus subtended 2° of visual angle. A schematic drawing of the apparatus is given in Figure 1.

Calibration

The success of the experiment was dependent on an accurate radiance measurement of the test stimulus. Because time and equipment limitations prevented a radiometric measurement of the test stimulus after each trial, an indirect calculation of the radiance was accomplished. This calculation was based on the concept that the test stimulus intensity equals the maximum test channel radiance for the stimulus wavelength multiplied by the transmission of the variable density filter system setting. The maximum test channel radiance was outside the range of both the lab's spectroradiometer and optical power meter. However, since this experiment was concerned with measuring relative luminous efficiency a relative measure of test channel radiance across wavelengths was sufficient. Relative spectral output of the test channel across wavelengths was measured with an iris aperture in place to reduce test channel radiance to measurable levels. The measurement of the relative spectral output was conducted

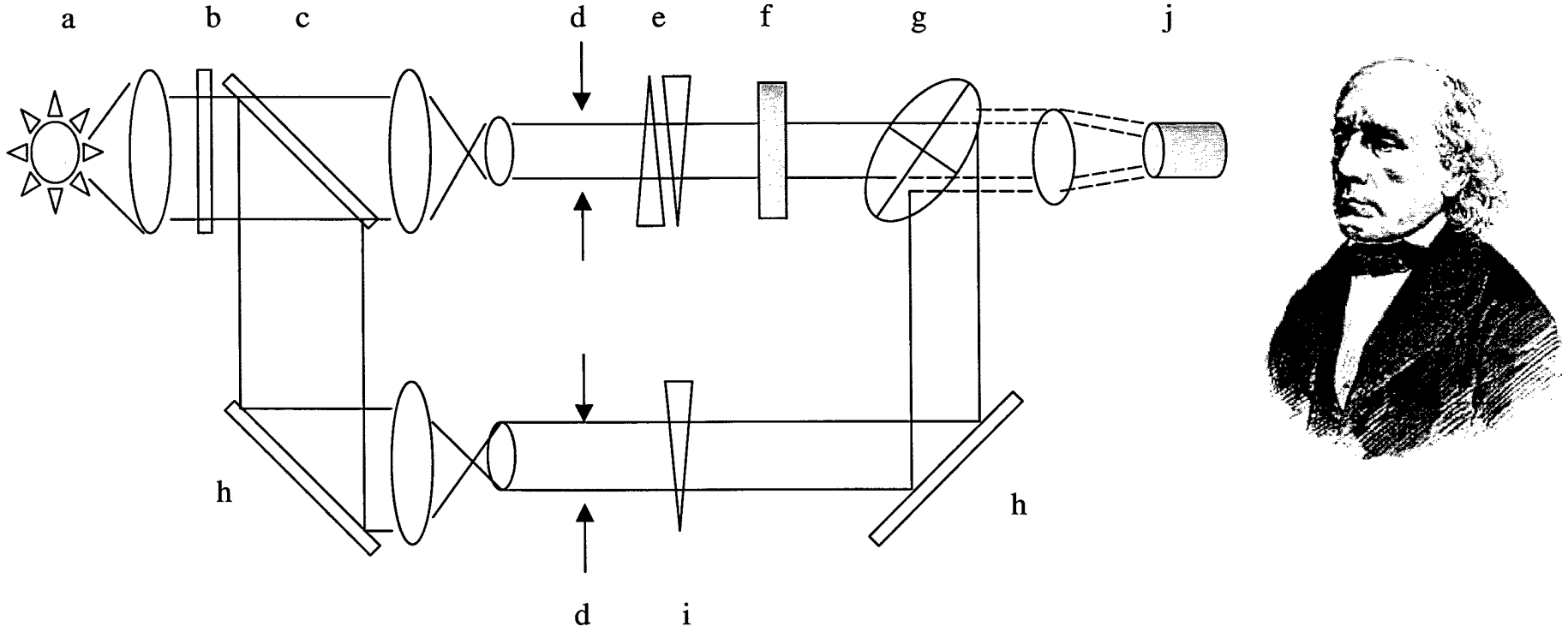

Figure 1: Two channel optical system: top channel produces the monochromatic test stimulus, bottom channel produces the achromatic reference stimulus. Components include a) 1000W Xenon arc lamp, b) heat absorbing lens, c) high transmission (99/1) beam splitter, d) variable aperture, e) counter rotating variable density filter system, f) monochromatic interference filter, g) rotating disc optical chopper with mirrored blades h) mirror, i) spectrally neutral, variable density filter, j) diffusing optic/backlit viewing screen

both before and after the experiment (Figure 2A) to verify that expected changes in spectral output (due to aging of the lamp and other factors) were small compared to the measured effect size. The lamp output shift slightly towards longer wavelengths, which suggests that there may be a small underestimate, in short wavelength sensitivity in the luminous efficiency curves. Figure 2B shows the impact of the lamp output shift on the 100fl heterochromatic brightness matching curve. The maximum change in calculated relative sensitivity between the pre-experiment and post-experiment calibrations for the 100fl heterochromatic brightness matching curve is .08 at 532nm.

Measurement of the test channel radiance and calibration of the variable density filter system was accomplished using a new, NIST-traceable, factory calibrated, Photo Research PR 650 spectroradiometer purchased for this project. Spectral calibration for the test stimulus channel was verified by cross checking the manufacturer's nominal filter transmission peaks and band widths for each of the narrow bandpass filters against the measured transmission peaks using the PR 650 Spectrascan spectroradiometer. Nominal specifications and measured characteristics were found to be in agreement for all seventeen filters.

The spectral transmission curves for the variable density filter system were found to vary with optical density. Therefore optical density calibration for the variable density filter system was accomplished for each of the 17 wavelengths separately. For each test wavelength, the optical density of the variable density system was sampled at 22-degree intervals using the Photo Research PR 650 spectroradiometer. These sample points were imported into Sigma Plot 5.0 and a non-linear regression was performed to model the

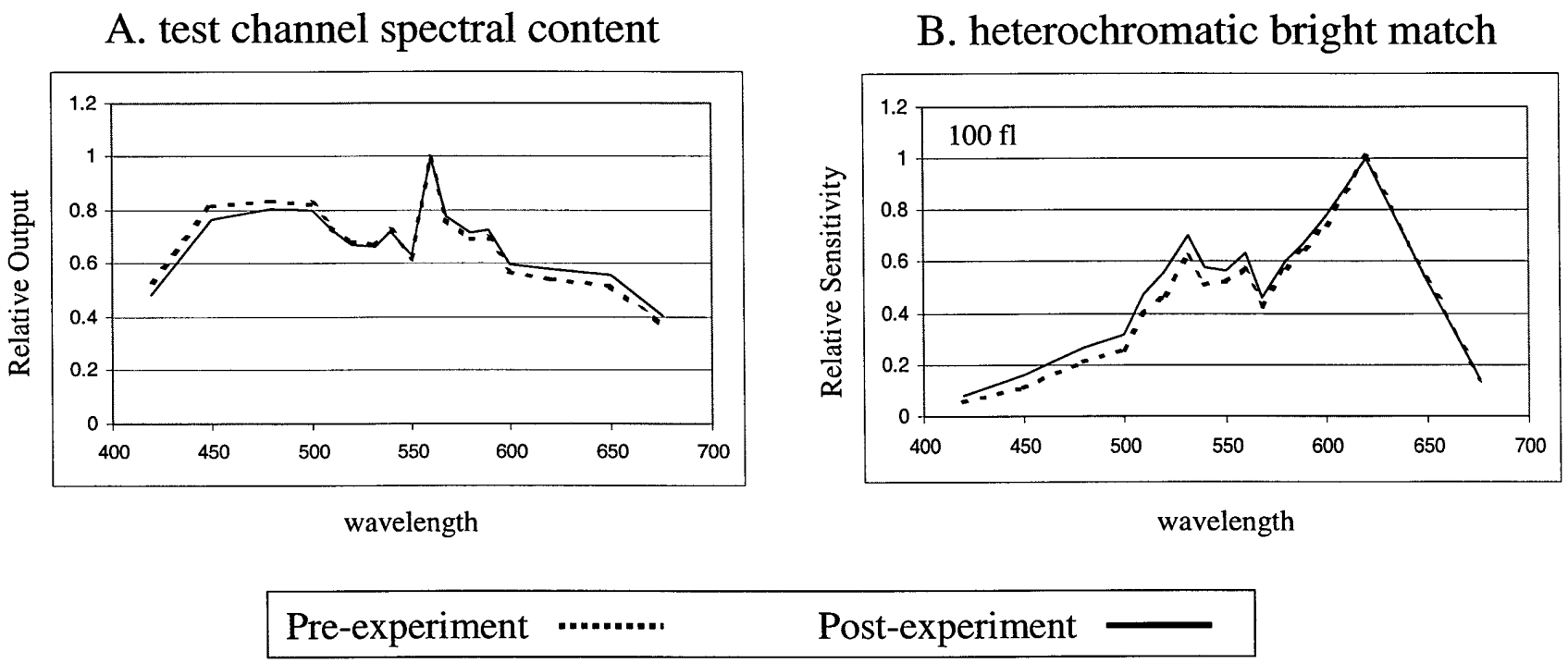

Figure 2. A: Change in spectral output of the test channel over the duration of the experiment due to changes in the optical system such as aging of the lamp. B: Calculated relative sensitivity under the 100fl heterochromatic brightness matching condition using the pre-experiment and post-experiment system spectral output. The error introduced by unavoidable changes in system spectral content is small compared to the effect size due to the independent variables of the experiment.

angular rotation vs. filter transmission function. The resulting regression equations fit the data points very well (r^2 >.98 across all wavelengths). The regression equations were used to calculate the transmission of the variable density filter system for each subject's flicker or heterochromatic brightness matches.

Adaptation

It is well known that maintaining an appropriate level of adaptation is critical to the success of any vision experiment. A common strategy for vision research is to allow subjects to dark-adapt before starting an experimental session. The advantage of dark adaptation is that the adapting stimulus is spectrally flat thus avoiding asymmetric fatigue of the visual system. Dark adaptation was not appropriate for this project for two reasons. First, a sudden change in luminance levels from near zero to 1000 foot lamberts (fl) would be uncomfortable and interfere with the quality of the measurement. Second, the luminance level during the trial would substantially raise the subject's adaptation state, thus requiring a large amount of time between trials to reestablish dark adaptation. The importance of maintaining appropriate adaptation for high intensity stimuli experiments was demonstrated during the Weber fraction experiments done by Alpern, Rushton and Torii (1970). In their investigation of increment thresholds on flashed backgrounds they found increasing Weber fractions at high luminance levels, a finding expected as cone saturation is approached. However if the subject was allowed to adapt to the background, the Weber fraction is constant across luminance intensities, even at intensities that approach pathological levels (Vos et al, 1972).

In this experiment, using the reference stimuli as the adapting source would have been analogous to the steady adapting source used by Alpern, Rushton and Torii, but the time it took to start and stop the chopper made this approach impractical. However, it was possible to rapidly configure the test channel to produce a white stimulus of equal luminance to the reference stimulus. Therefore, during the 30 second adaptation period prior to running the set of trials for each wavelength our subjects viewed the alternating, equal luminance, broad spectrum, stimuli produced by the reference and test channels. Immediately after adaptation, the appropriate monochromatic interference filter was rotated into the test channel and the matching trials began. The adapting stimuli were spectrally broad but not spectrally flat. In addition, the spectral content of the adapting stimuli varied with reference intensity. The correlated color temperature for the adapting sources varied between 3229 K and 5532 K as shown in Figure 3.

Procedure

The xenon lamp was allowed warm up for a minimum of 30 minutes to reach equilibrium and then the reference channel was adjusted to one of four luminance levels (1, 10, 100 or 1000 foot-lamberts) as measured by the Photo Research PR 650. Next the test channel's interference filter wheel was rotated to the open (achromatic) position and the test channel's luminance was adjusted either using the iris aperture (1 and 10fl levels) or the variable density filter system (100 and 1000fl levels) to match the luminance of the reference channel. Then the chopper was started at either 2 Hz for the heterochromatic brightness matching trials or 20-30Hz for the flicker photometry trials. The alternating achromatic, equal luminance, reference and test fields served as an adapting source,

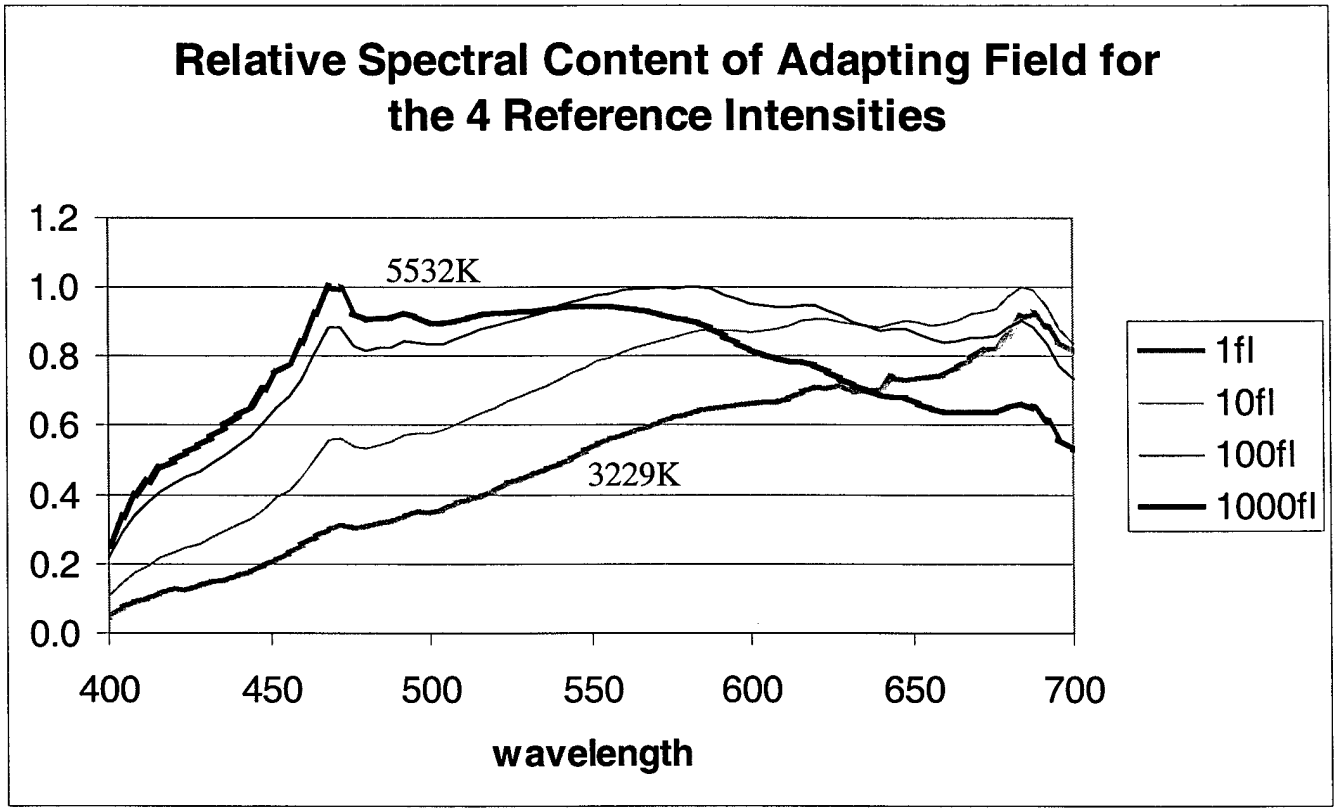

Fig 3: Spectral content of the adapting field for the four levels of reference intensities. The adapting field was a broad-spectrum, white, field for all conditions, however the blue content of the adapting field increased at the higher reference intensities. The correlated color temperature for the adapting sources varied between 3229 K(1fl) and 5532 K (1000fl)

viewed by the subjects, before each set of trials. After the subject viewed the adapting field for 30 seconds, one of the wavelength interference filters, selected randomly without replacement, was rotated into the test channel. The subject then adjusted the test field brightness to match the reference field using the variable density filter system. Once the match was made, the experimenter recorded the variable density filter system setting in a spreadsheet. The experimenter then decreased the test channel intensity by at least 2 log units and the subject made another match. Four trials (matches) were made in immediate succession. Each trial began with the test channel intensity set at least 2 log units below the match intensity.

After completing the four trials for a given wavelength the subject viewed the adapting source for 30 seconds and a set of trials were started at the next wavelength in the random sequence. This process was repeated until all 17-test field wavelengths had been completed.

Data Reduction

The variable density filter system settings from each trial were converted to system transmission values using the regression equations derived from the filter calibration. Then the 4 transmission values for each wavelength were averaged. These average transmissions were multiplied by the respective test channel radiance value to produce an averaged test stimulus radiance value for each wavelength. The reciprocal of these "equal brightness" radiance values produced the relative sensitivity values. A Z-score transformation was performed on the relative sensitivity values. Each subject's sensitivity scores were transformed using the mean and standard deviation for their own

scores at the given method and reference intensity. The resulting Z-scores were used to conduct a 3 way (method by intensity level by wavelength) repeated measures analysis of variance (ANOVA). Finally, a traditional luminous efficiency transformation of the sensitivity values was calculated. The relative sensitivity value at each wavelength was divided by the peak value, producing a curve with a maximum of one. These luminous efficiency values were plotted for each subject under each condition. The 3-way ANOVA was repeated with the calculated relative values of luminous efficiency.

Subjects

The institutional review board of the University of Missouri –St. Louis approved all of the experimental procedures used in this study including the informed consent form signed by each subject. Seven subjects, 4 females and 3 males, ranging in age between 22 and 41 years old, were selected to participate in this study. Subjects were screened for normal color vision using psuedoisochromatic plates and were paid for their participation. All 7 subjects completed all of the measurements. Prior to data collection, each subject practiced between 2 and 6 hours, an average of 3.1 hours, on the tasks. Most of the practice was dedicated to heterochromatic brightness matching. Practice was divided over several days with no practice session lasting more than one hour.

During heterochromatic brightness matching practice sessions, 2 of the subjects could make repeatable (range less than .5 log unit) matches using only the brightness of the stimuli in their decisions. The other 5 subjects had a much larger (>1 log unit) range over which the reference and test stimulus seemed to be of equal brightness. These subjects were encouraged to standardize their match procedure. However, even with

additional training and standardization of the procedure, several of the subjects felt they needed to incorporate additional criteria into their match decisions. One subject looked for a characteristic appearance to the halo (glare) surrounding the test stimulus. Another subject looked for chopper blade motion cues that could be seen in the stimuli when the test and reference where approximately equal intensities. These additional criteria were difficult for the subjects to describe and impossible to quantify. No effort was made to discourage the additional criteria as long as the subject reported that the criteria were present for all 17 wavelengths tested.

Results

The only physical difference between the flicker photometry task and the heterochromatic brightness matching task was the rate at which the reference and the test stimuli alternated. Still this small change in physical parameter had a profound impact on the subjects' performance including the match criteria, match reliability, training requirements and most importantly the luminous efficiency curve produced. It is worth describing the results unique to each task separately before looking at the overall analysis.

Flicker

Subjects found the flicker photometry task relatively easy to perform. After a brief explanation of the testing procedure, all subjects were able to make rapid and

repeatable flicker matches with no more than 15 minutes of practice. Subjects could complete a data collection session in approximately 1 hour. Figure 4A shows a typical distribution of the transmission values selected by one subject for the flicker photometry condition. In the flicker condition, intra-wavelength variability was small compared to the magnitude of change in transmission values between wavelengths.

Figure 5 plots the mean relative sensitivity settings of each subject at each of the 4 reference intensities. The scatter plot points are well ordered and the overall pattern to the data is easily recognized. The difference between the peak sensitivity of the least sensitive and most sensitive subjects was approximately a factor 2 at each level. Rank order of subject sensitivity at one level was not predictive of rank order at other levels. For example subject 3 was the least sensitive subject at the 100fl level but the most sensitive subject at the 1000fl level. Direct comparisons of absolute sensitivity between intensity levels cannot be made due to a difference in relative sensitivity scaling for each reference intensity level, which resulted from the limited measurement range of the spectrophotometer.

Relative changes in sensitivity across intensities were examined by converting the relative sensitivity values to Z-scores. The between-subject mean curve for these transformed scores is shown in Figure 6A. The curves are narrow and unimodal with peaks around 560nm. The overall shape of the curves are strikingly similar to $V(\lambda)$. The 1fl and 10fl curves are nearly identical but the higher intensity 100fl and 1000fl curves appear slightly narrower then the lower intensity curves.

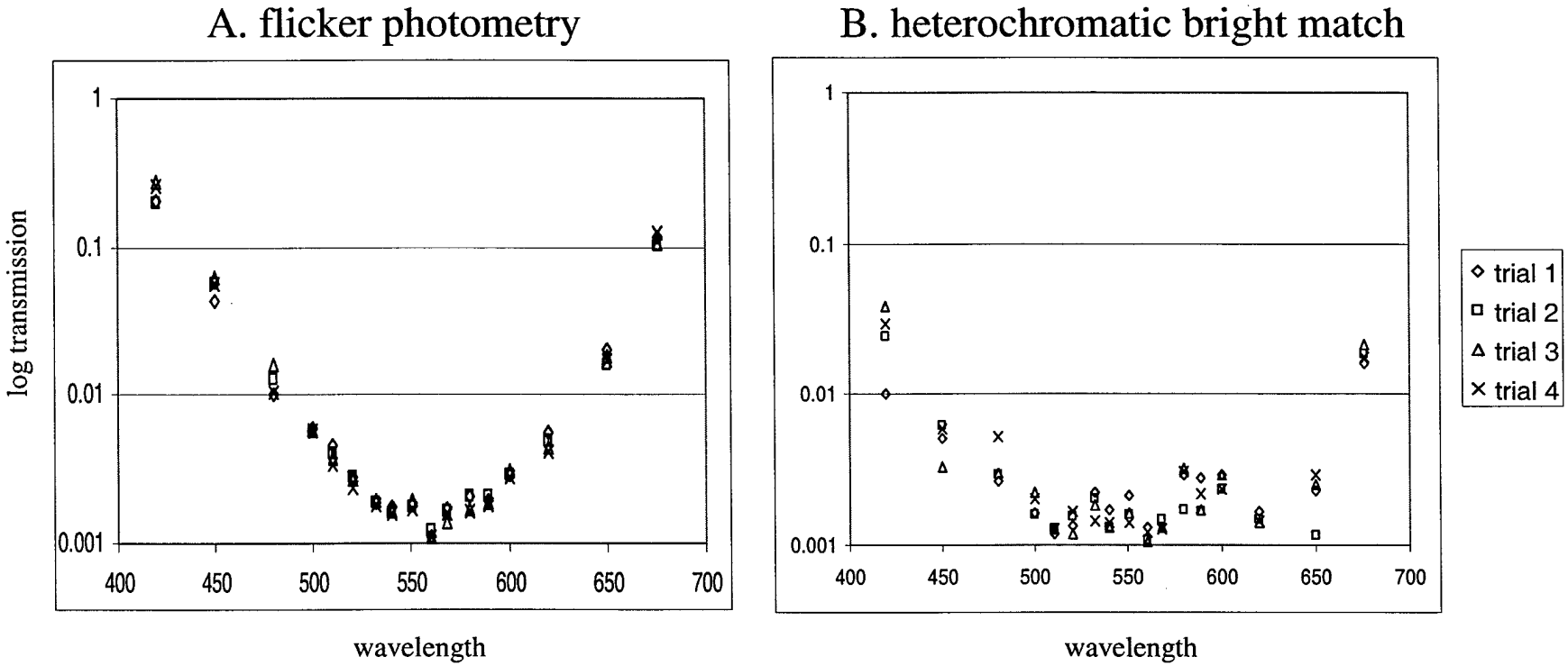

Figure 4, A and B: Typical transmission values of the test channel for the four trials at each wavelength for flicker photometry (A) and heterochromatic brightness matching (B) methods. Data from subject 4 was taken at 10fl reference intensity. Transmission values were less variable in the flicker condition for all subjects.

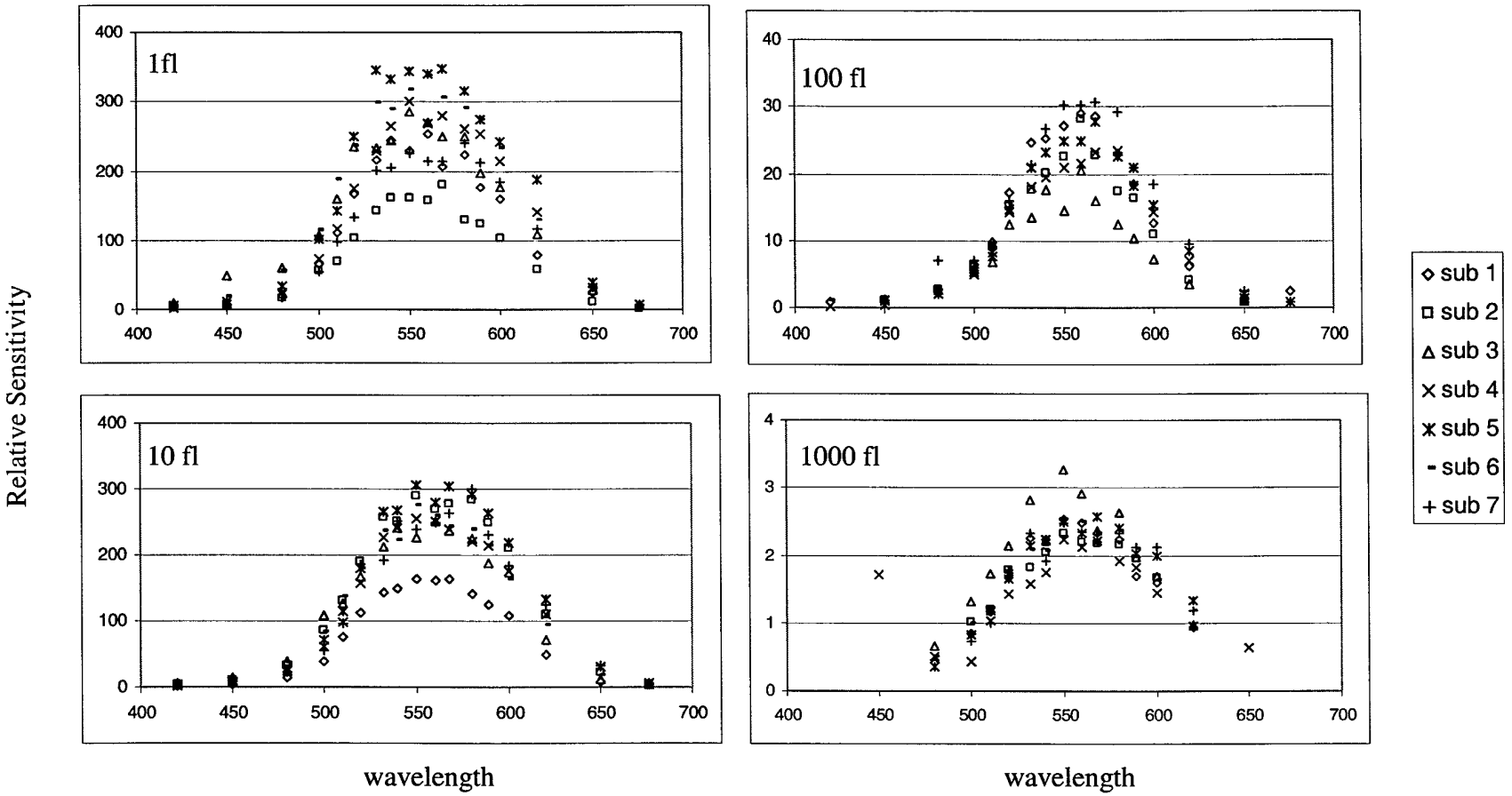

Figure 5: Relative sensitivity for all subjects by flicker photometry for the four reference level intensities

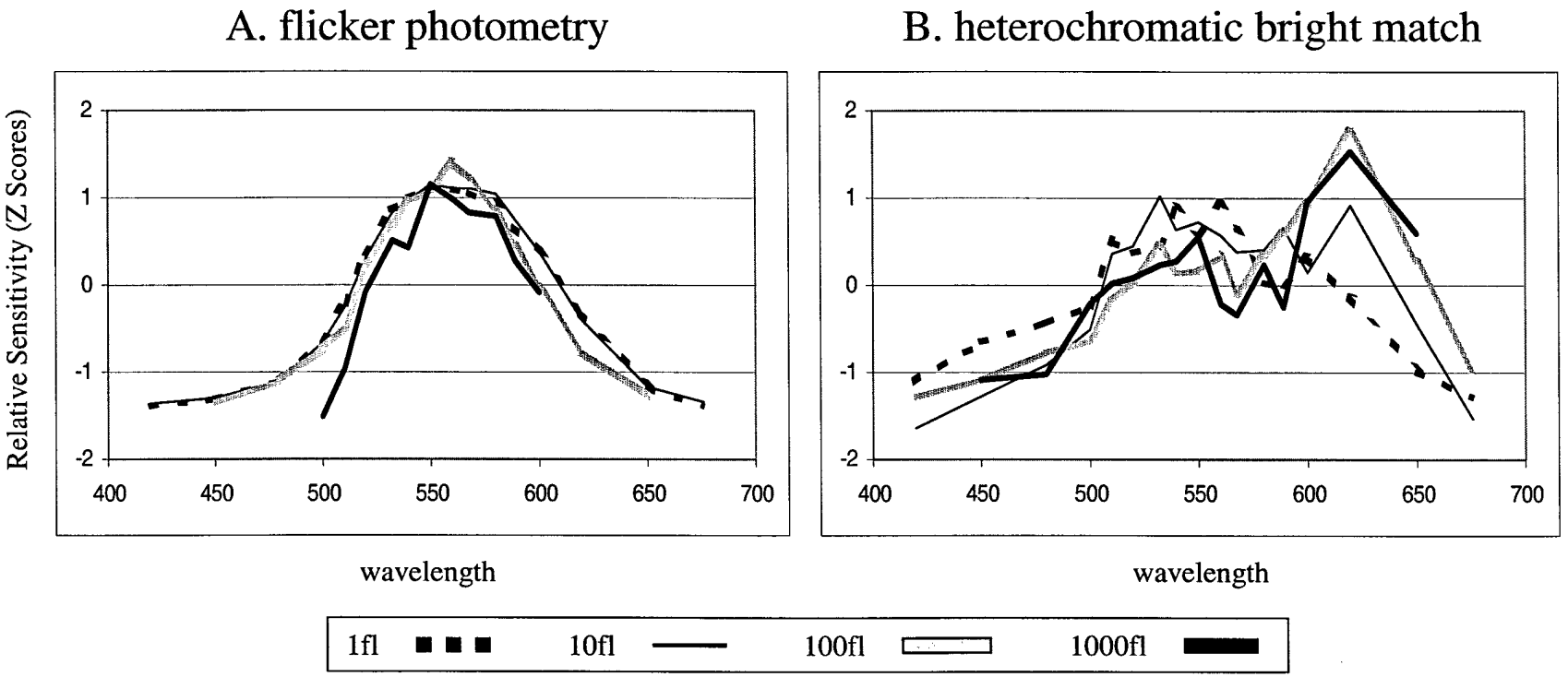

Figure 6, A and B: Mean relative sensitivity (Z scores) across subjects for each of the 4 levels of reference intensity. The flicker photometry curves (A) have the classic Vλ appearance, narrow with a single peak around 555nm, and are consistent across the reference intensities. The heterochromatic brightness matching curves (B) have a broader appearance with multiple peaks. Maximum sensitivity shifts with reference intensity from 560nm for the 1fl level to approximately 620nm for the higher intensities.

Successive Heterochromatic Brightness Matching

All subjects found the heterochromatic brightness matching task to be considerably more difficult than the flicker photometry task with each data collection session lasting approximately 90 minutes. Despite extensive practice, variability in each subject's four settings at each wavelength remained large. One example of this variability is shown in Figure 4B. The variability in heterochromatic brightness matching was generally larger than the variability on the flicker task for all subjects. There was also more variability relative to the magnitude of change across wavelengths. The later indicates a lower signal to noise ratio for heterochromatic brightness matching.

The mean settings of each subject at each wavelength for the four reference intensities are shown in Figure 7. In the 1fl condition, the overall shape of the points is similar in appearance to the corresponding flicker data (Figure 5); however, the difference between the data of the most sensitive and least sensitive subject was much greater than in flicker photometry condition. There was at least a factor of 10 difference for the heterochromatic brightness matching data while the difference was only a factor of 2 for flicker photometry. The patterns in the data points for 10fl, 100fl and 1000fl charts in Figure 7, are more difficult to discern than they are for the corresponding flicker levels. However, the rank order in heterochromatic brightness matching sensitivity is better preserved across intensity levels than it was in the flicker condition. Subjects 5 and 6 are more sensitive and subject 3 is less sensitive for most wavelengths at all levels of intensity.

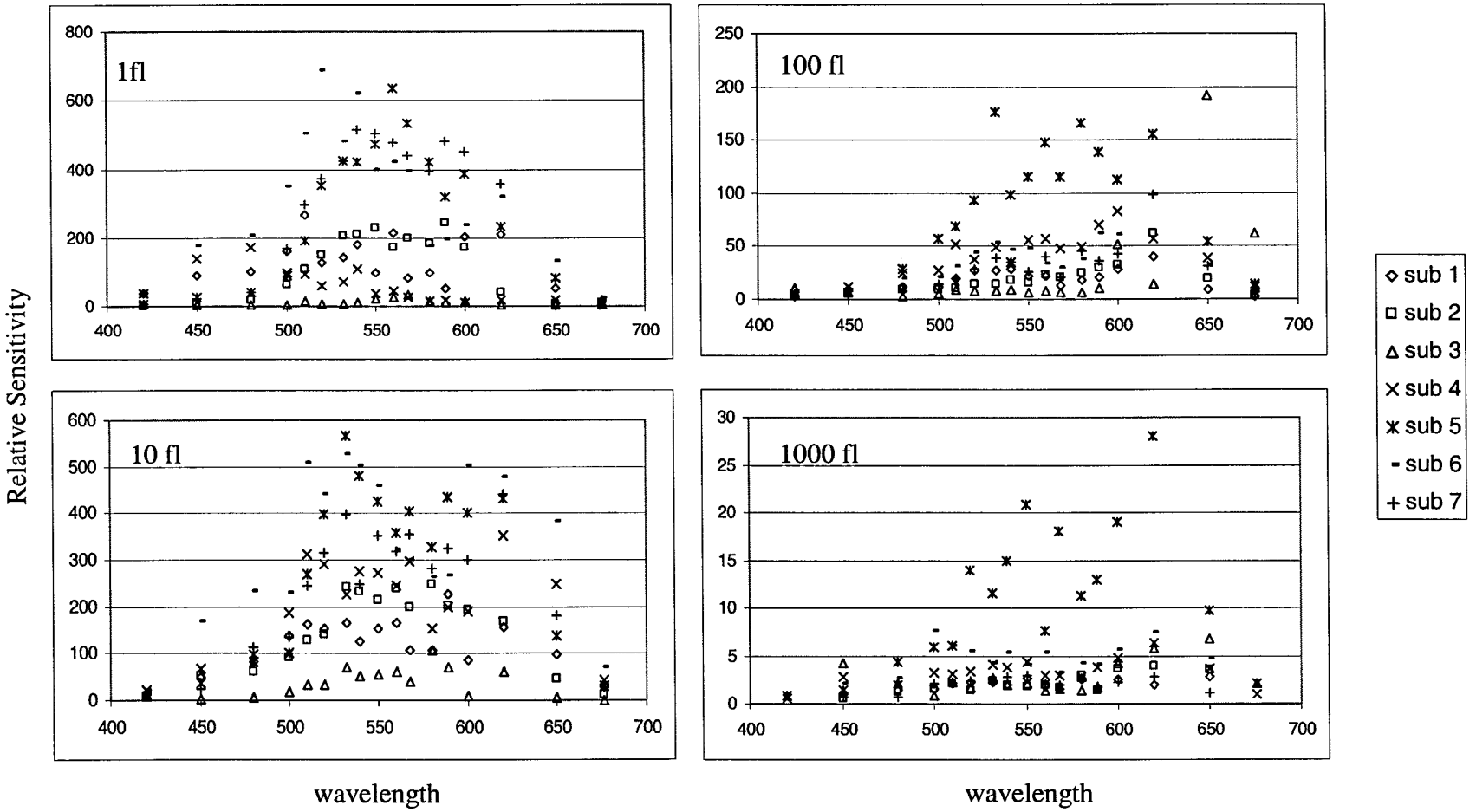

Figure 7: Relative sensitivity for all subjects by heterochromatic brightness matching for the four reference level intensities

The between-subject mean of the Z-transformed sensitivity scores are plotted for each reference intensity in Figure 6B. The mean heterochromatic brightness matching curves are broad and show multiple peaks. For all four intensities, there is a middle wavelength peak near 550nm, a notch between 550nm and 600nm, and a long wavelength peak at 600nm or higher. With increasing intensity, the size of the longer wavelength peak grows relative to the middle peak and the wavelength of maximum sensitivity shifts with from 560nm to approximately 620nm. The breadth of the curves increases with increasing intensity, particularly in the long wavelength portion of the spectrum.

Combined analysis

The comparison of relative sensitivity data across methods and intensities is complicated by the need to normalize the data. A common approach is to equate all curves at one wavelength (Wagner and Boynton, 1972; Abramov and Gordon, 1977; Yaguchi and Ikeda, 1980; Ikeda et al, 1982; Ikeda and Nakano, 1986). In many data sets this approach would be useful; however, in our case this procedure gave an inappropriate weighting to one subject (# 3) who had extremely low sensitivity values for all wavelengths except for 650nm in the 100fl heterochromatic brightness matching condition. When we averaged the data across subjects using this technique, the maximum sensitivity occurred at the 650nm wavelength. This representation of the data seemed distorted considering the relative sensitivity of the other 6 subjects at that wavelength as shown in Figures 8-14.

Subject 1

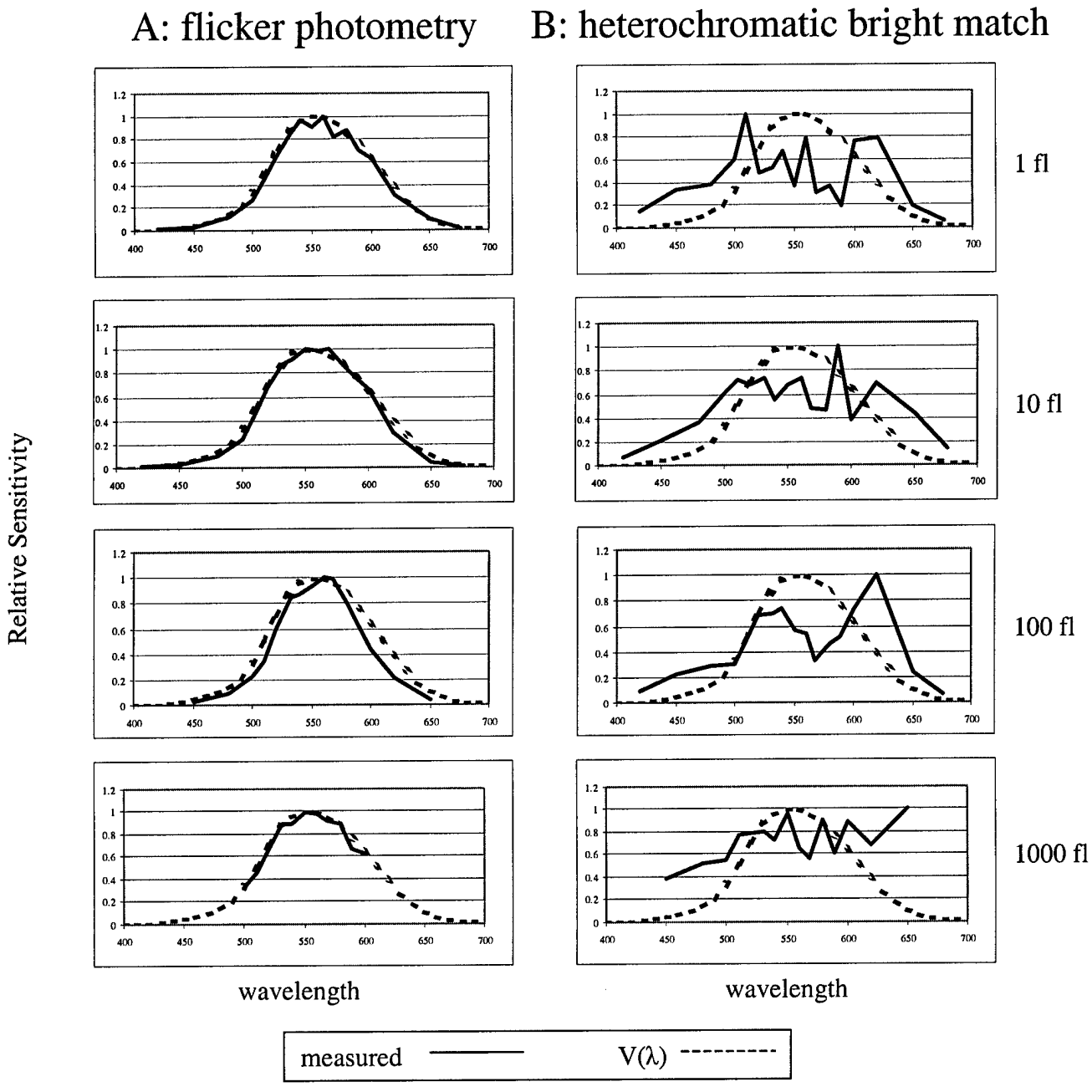

Figure 8: Luminous efficiency curves for subject 1 for all conditions

Subject 2

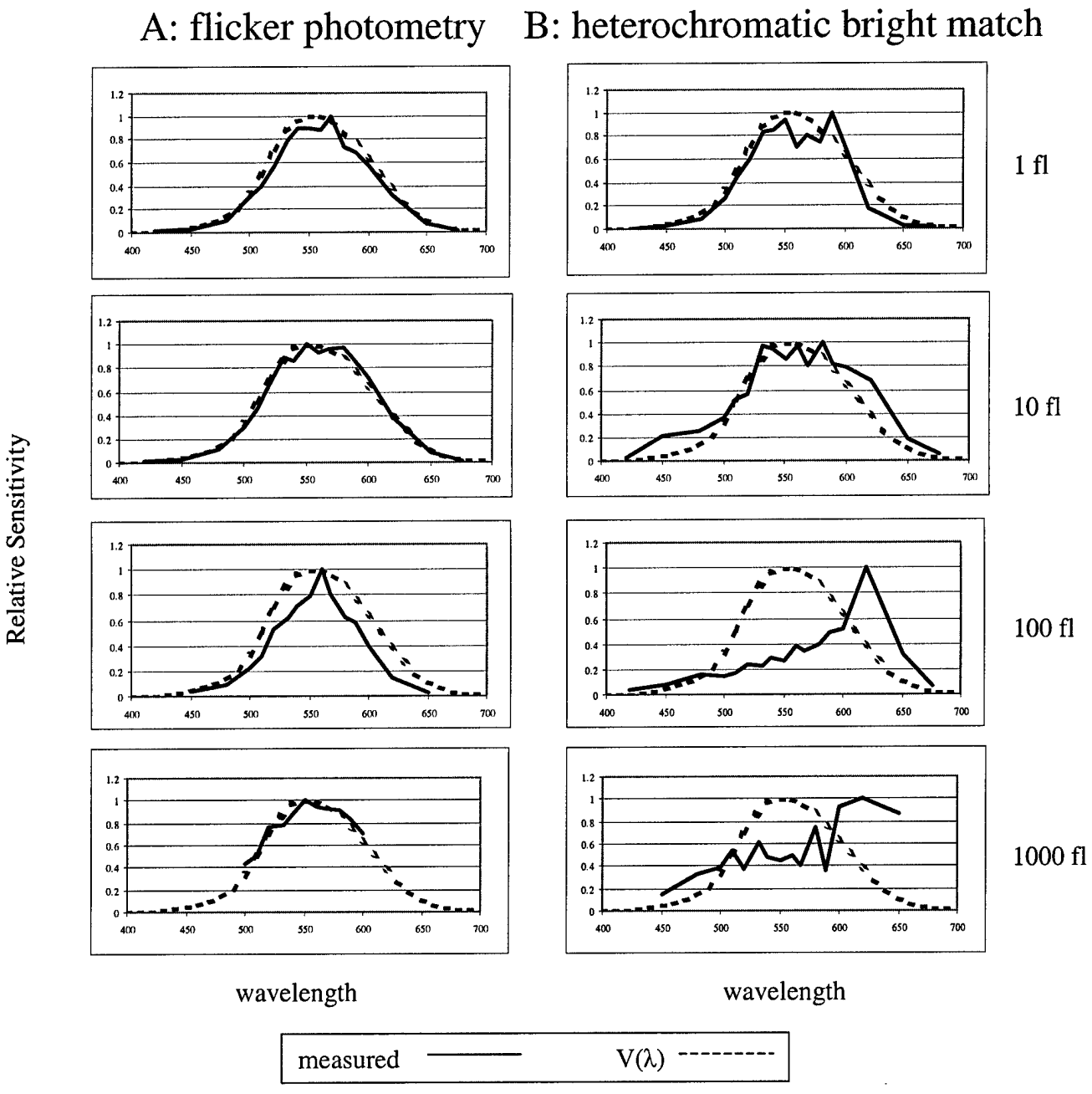

Figure 9: Luminous efficiency curves for subject 2 for all conditions

Subject 3

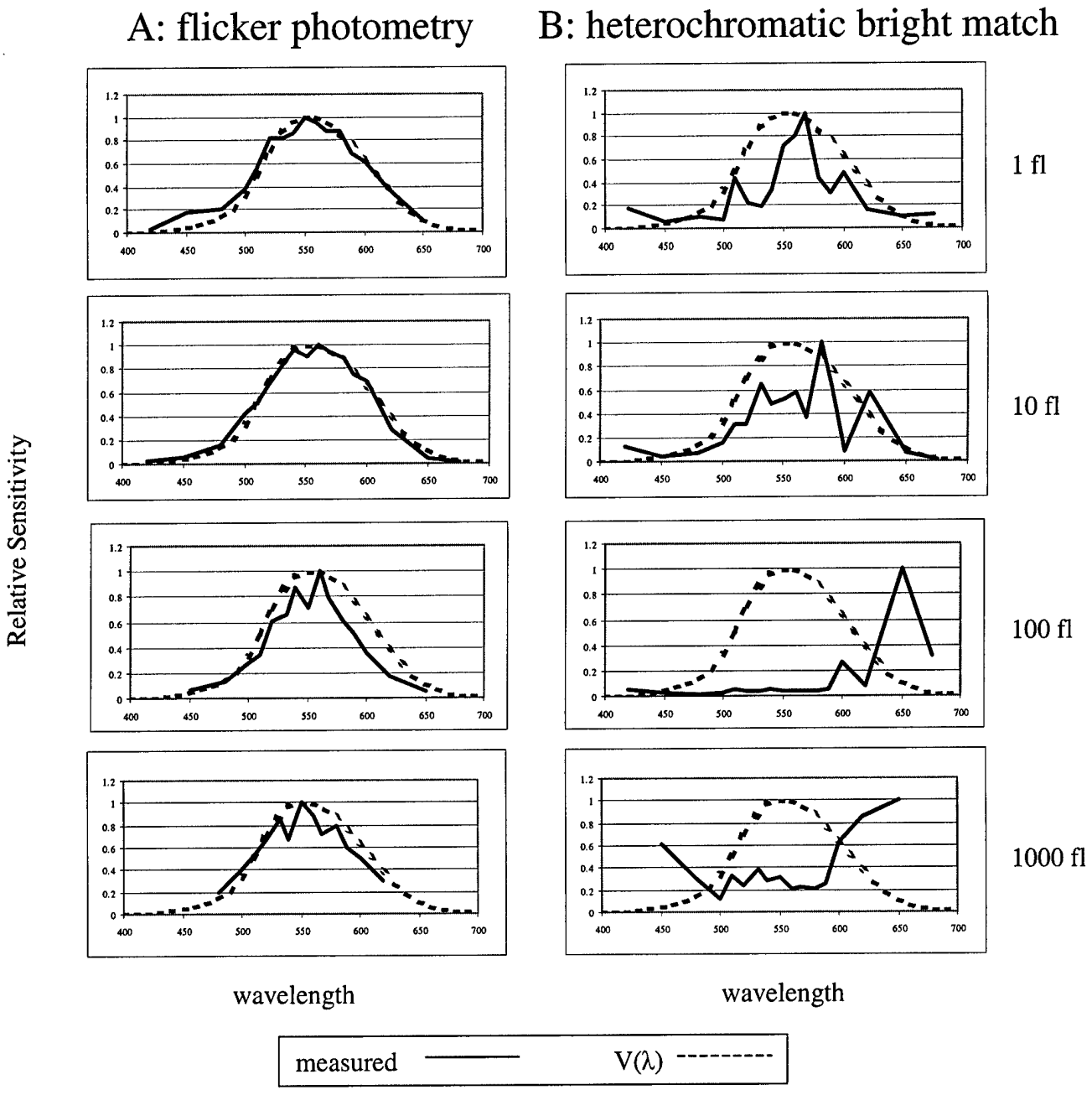

Figure 10: Luminous efficiency curves for subject 3 for all conditions

Subject 4

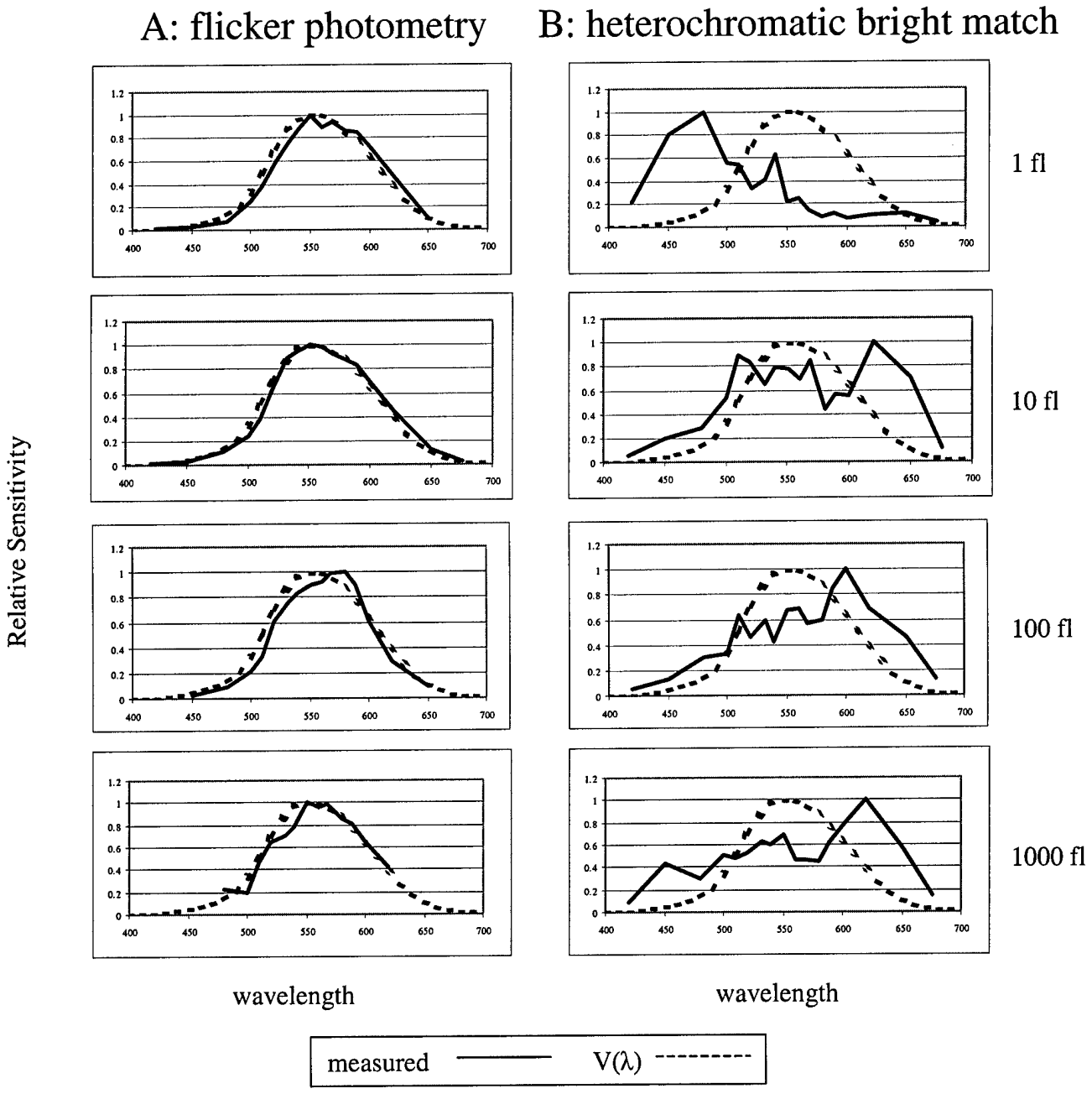

Figure 11: Luminous efficiency curves for subject 4 for all conditions

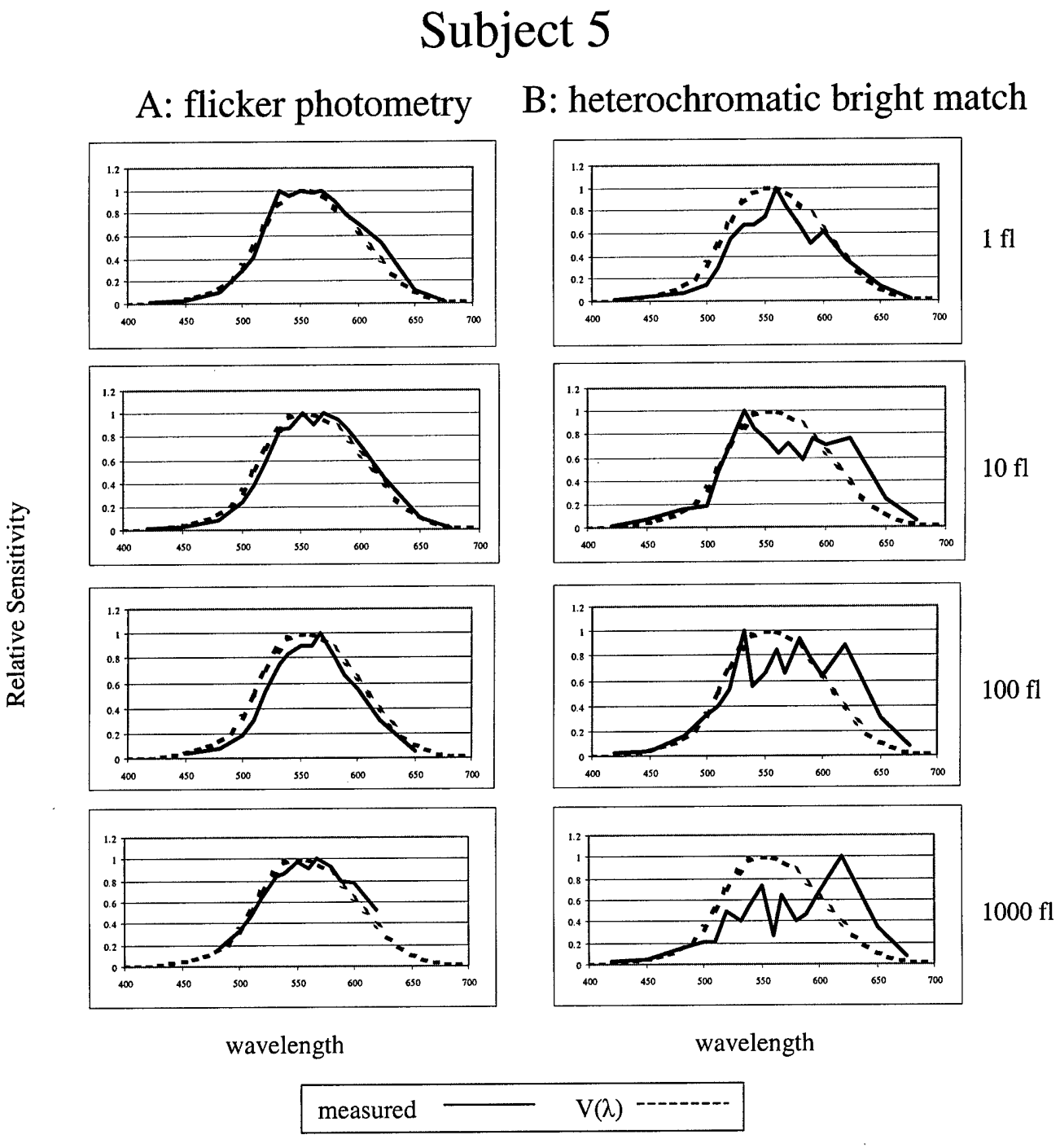

Figure 12: Luminous efficiency curves for subject 5 for all conditions

Subject 6

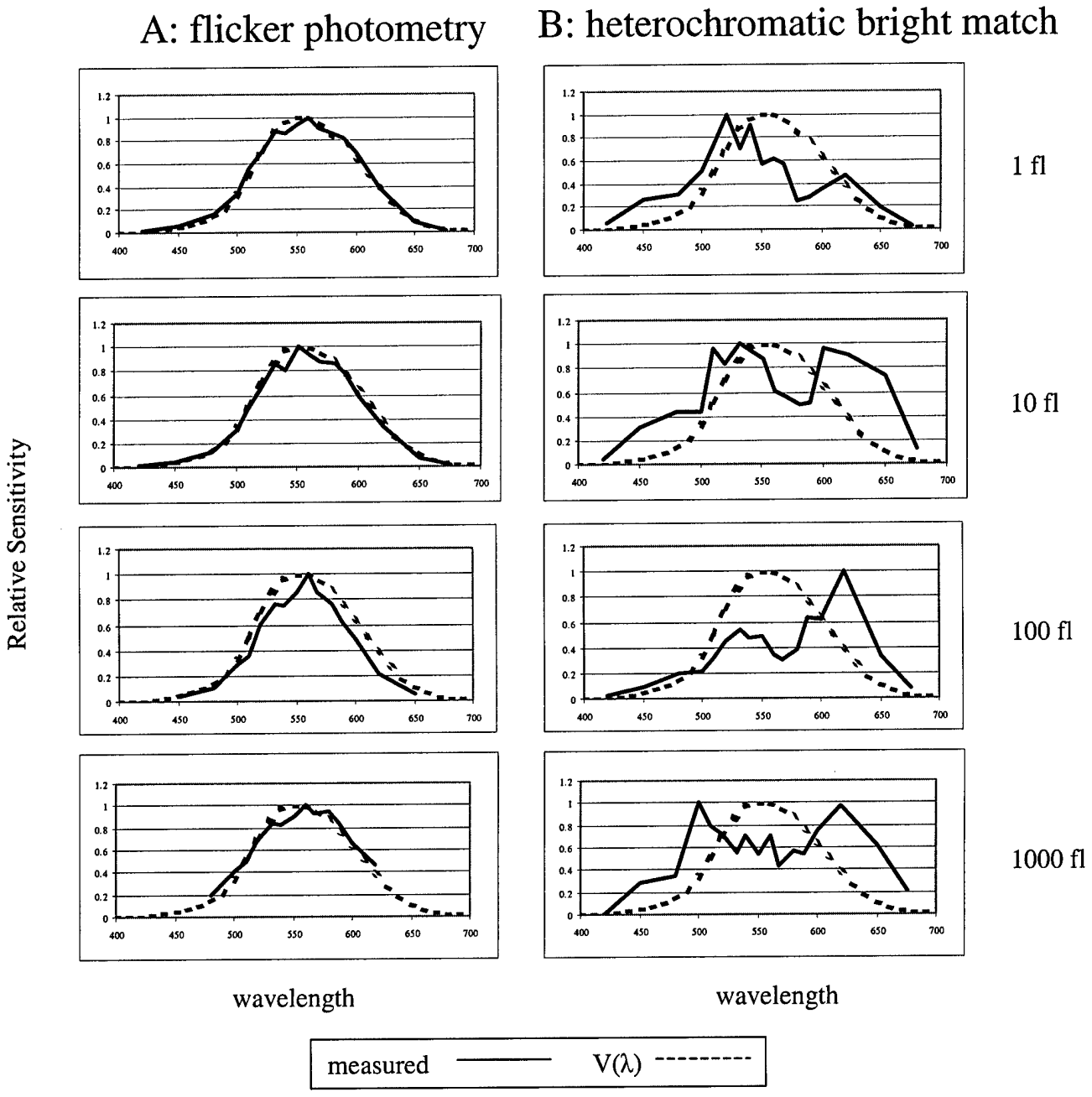

Figure 13: Luminous efficiency curves for subject 6 for all conditions

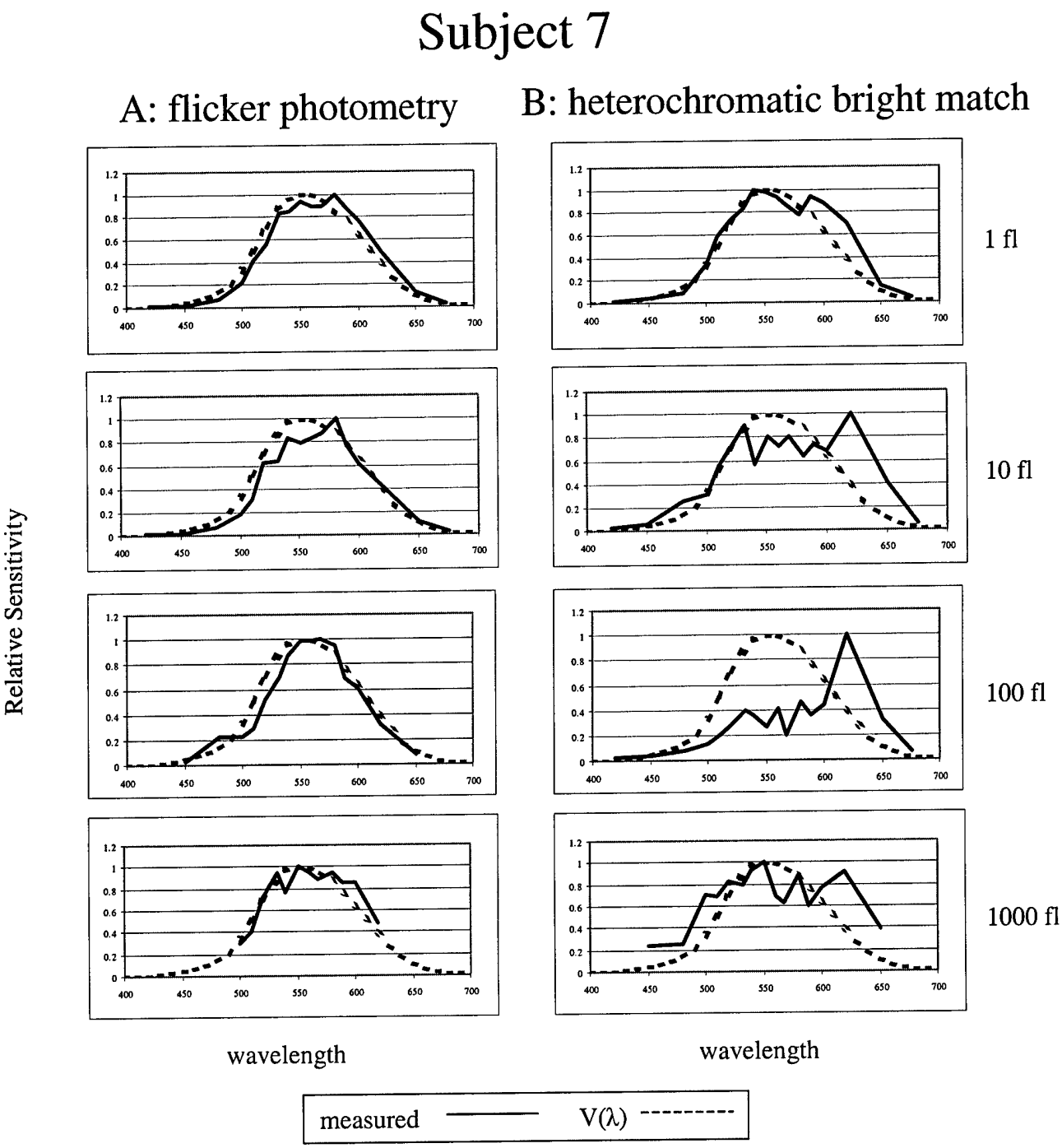

Figure 14: Luminous efficiency curves for subject 7 for all conditions

Because of the difficulty with this data transformation, I decided to use a Z-score transformation. Each subject's sensitivity scores were transformed using the mean and standard deviation for their own scores for the given method and reference intensity. This transformation, which brings the mean for each subject's sensitivity curve to zero and set the standard deviation for each curve at one, permits the direct comparison of relative sensitivity across reference intensity levels.

I conducted a Method (Flicker Photometry, Heterochromatic Brightness Match) by Level (1fl, 10fl, 100fl, 1000fl) by Wavelength (500, 510, 520, 532, 540, 550, 560, 568, 580, 589, 600) analysis of variance (ANOVA) on these Z-scores. Since ANOVA uses variance to look for mean differences and since the Z-transformation equates the means to zero and the standard deviations to one, by definition this analysis could not show any main effects for Method or Level. However I retained the ability to examine the main effect of Wavelength and all interactions except for Method by Level. The outcome for this analysis was significant for Wavelength ($p<.001$), Method by Wavelength ($p<.001$), Level by Wavelength ($p<.001$) and the three way interaction Method by Level by Wavelength ($p<.001$). The main effect of Wavelength simply confirms that the eye is not equally sensitive to all wavelengths of light. The significant interactions of Method by Wavelength and Level by Wavelength support the conclusion that the shape of the sensitivity curve changes with the method for measuring sensitivity and the intensity of the reference. More specifically, as seen in Figure 6B, with heterochromatic brightness matching, the peak relative sensitivity of the eye changes from a green wavelength around 560nm into the red (between 600 and 650nm) as

intensities increased. It is obviously desirable to compare this change in relative sensitivity at higher intensities with V(λ). Unfortunately there is no direct way to transform the Z-score or V(λ) to bring them to the same scale.

To directly compare this study's results with V(λ) a different data transformation is required. V(λ) is scaled so the maximum sensitivity is 1. A relative sensitivity curve was calculated for each subject under each condition. For each of these 56 curves (7 subjects times 8 conditions) the relative sensitivity values were divided by the maximum sensitivity for that subject and condition. With each curve scaled, the individual data can be compared directly to V(λ) (Figure 8-14). An ANOVA of the transformed relative sensitivity was performed. In this transformation the means and standard deviations for the curves are brought close together but not equated. This permitted the testing of the main effects for Method and Level, which had been negated in the Z-score transformation.

The ANOVA was based on data collected at the wavelengths between 500 and 600nm; more extreme wavelengths were excluded because of missing values. There were significant main effects of Method ($p = .004$) and Level ($p = .008$). There were also significant effects for Wavelength, Method by Wavelength, Level by Wavelength and the three way interaction, Method by Level by Wavelength ($p < .001$). These results were similar to those obtained from the Z-score analysis.

The between-subject mean curve of the V(λ) style transformation scores are plotted in Figure 15. In this figure, the mean relative sensitivity curves for the two methods, heterochromatic brightness matching and flicker are compared to V(λ) at each

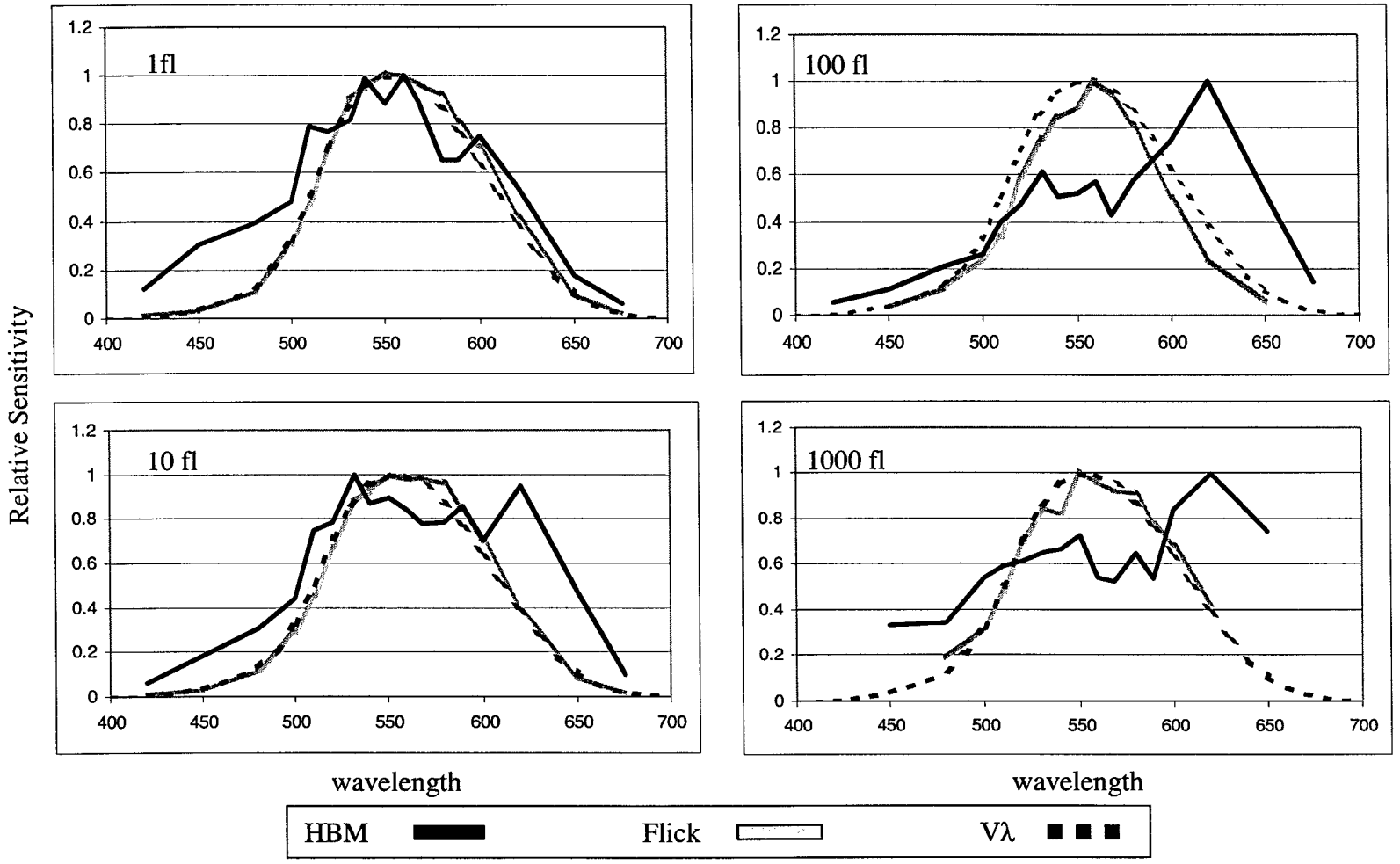

Figure 15: Average luminous efficiency for the seven subjects by heterochromatic brightness matching (HBM) and flicker photometry (Flick). The standard CIE Luminous Efficiency for Photopic Vision (Vλ) is included for comparison.

intensity level. This graph shows that the flicker curves are very similar to V(λ) in shape, though the width of the 100fl flicker curve is narrower than V(λ) and the other flicker curves. The heterochromatic brightness matching curves are broader then V(λ) and the flicker curves at all intensities and the peak of the curve migrates towards 620nm with increasing intensity while the flicker peak remains constant. It is possible to visualize the main effect of Method in this figure. Peak sensitivity for heterochromatic brightness matching at the 100fl and 1000fl levels is outside of the 500 to 600nm range for which the ANOVA was run causing the mean for heterochromatic brightness matching (.560) to be below the mean for flicker (.734).

Visualizing the main effect of Level in Figure 15 is more challenging than it is for Method. Both the 100fl heterochromatic brightness matching curves and the 100fl flicker curves seem to be lower than their respective curves at the other intensities. Since the main effect for Method was significant, we analyzed the Level factor separately for the flicker and heterochromatic brightness matching conditions. Under both conditions, the effect of Level was found to be significant (for flicker $p = .002$ and for heterochromatic brightness matching $p = .037$). The pairwise comparison for the flicker data showed that the relative sensitivity was lower for the 100fl level than for 1fl ($p = .008$), 10fl ($p = .025$) and 1000fl ($p = .007$). Pairwise comparisons for heterochromatic brightness matching found that relative sensitivity was lower at the 100fl level than at the 10fl level ($p = .013$).

Figure 16 shows the flicker graphs from Figure 5 along side the heterochromatic brightness matching graphs from Figure 7 for the 1fl and 1000fl levels. The scaling has

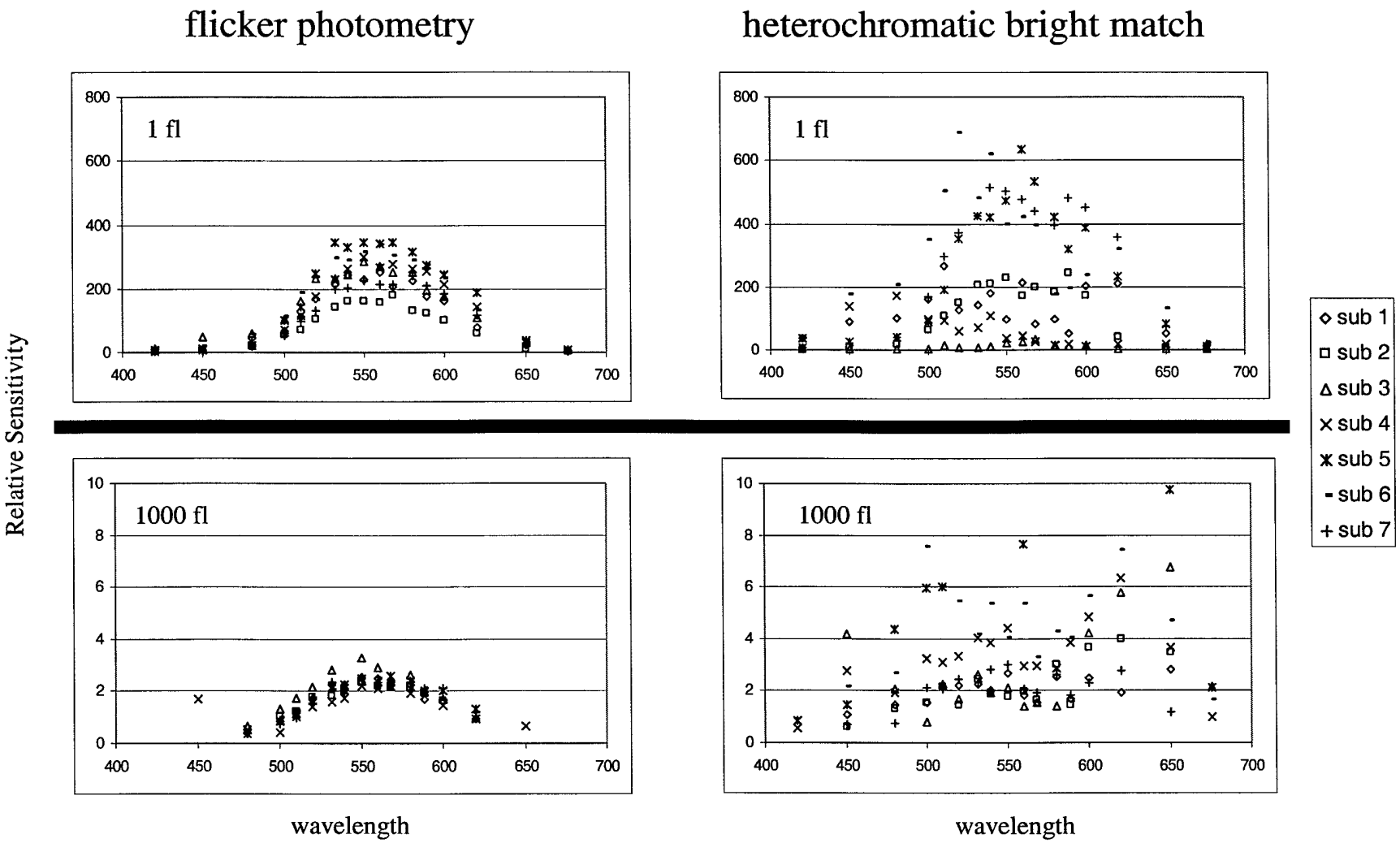

Figure 16: The flicker charts from Figure 5 are shown along side the heterochromatic brightness matching charts from Figure 7 for the 1fl (top)and 1000fl (bottom) intensities. The scaling has been equated on the ordinate so that a direct comparison of sensitivity can be made between the two conditions.

been equated on the ordinate so that a direct comparison of sensitivity can be made between the two conditions. For the 1fl graphs on the top of Figure 16, the heterochromatic brightness matching points are more variable than the flicker points at each wavelength but there is no obvious difference in overall sensitivity between the two methods. A two-tailed paired t-test showed no statistical evidence for a difference in sensitivity ($p = .09$). However, for the 1000fl charts on bottom, the absolute sensitivity was higher for heterochromatic brightness matching than for flicker for most subjects and at most wavelengths. The two-tailed paired t-test at this intensity was highly significant ($p = 4.9X10^{-7}$).

Discussion

The major conclusion drawn from this study is that the shape of the heterochromatic brightness matching relative sensitivity curves depend on reference stimulus intensity. As reference intensity is increased, the relative sensitivity in the long wavelength region of the visible spectrum increases, which results in the peak sensitivity of the curve shifting from approximately 540nm at the 1fl intensity to over 600nm for intensities equal to or greater than 100fl. This shift in long wavelength sensitivity is accompanied by an overall increase in heterochromatic brightness matching sensitivity relative to that found by the flicker method. The shapes of the flicker photometry curves were also found to be dependent on reference intensity, although this effect was much less dramatic than in the heterochromatic brightness matching condition. The flicker curve was found to be slightly narrower for the 100fl intensity. In contrast to the

brightness matching, the flicker wavelength of peak sensitivity remained constant at approximately 550nm for all levels.

The interpretation of these results should be conducted within the context of the challenges encountered during data collection. The majority of these challenges are derived from the large intra and inter-subject variability routinely found in psychophysical measurements. It is common in investigations of luminosity to have subjects and data sets that are difficult to work with. In their landmark study Gibson and Tyndall (1923) noted that:

> An examination of the original data immediately shows that some observers were able to duplicate their ratio values very closely on different days, but others could not do so well.

The subjects with poor repeatability created difficulties for Gibson and Tyndall's analysis. They could not definitively separate those subjects who were simply poor at making photometric matches from others who had genuine fluctuations in their relative sensitivity. Not wanting to eliminate the variable subjects completely, Gibson and Tyndall decided to divide their subjects into good and poor repeatability groups, and then gave the good group twice as much weight in their computations.

Some of my subjects were also more consistent than others in their matches. This lack of consistency could be due to a variety of factors including instability in match criteria and differences in adaptive state.

Instability in match criteria over extended periods of time is more of a problem for heterochromatic brightness matching than for flicker photometry (Wagner and Boynton 1972). On one instance, during data collection, an equipment malfunction

interrupted a brightness matching session for several hours. During that break the subject's criteria shifted and we were unable to continue with the session where we left off. Heterochromatic brightness matching criteria were relatively consistent for data collected in one continuous session. This was verified by checking sensitivity against the order of testing on curves where there were dramatic changes in sensitivity for adjacent wavelengths, such as for subject #3 for 10fl heterochromatic brightness matching condition (Figure 10B). No systematic order effects were found on these inspections.

One possible threat to heterochromatic brightness matching criteria stability came from the introduction of experimenter bias. The subjects were very susceptible to suggestion. For example, simply having subjects manipulate the variable density filter system controller could cause dramatic changes in brightness perception, even when the filter wheel was disengaged. In addition if the experimenter made any comment about a subject's performance their heterochromatic brightness match could move up or down a log unit in intensity. Because of these observations, I took three steps to reduce the influence of the investigator on the subject's performance. First, once the subjects were trained, all the experimental sessions were run without giving the subject any specific feedback as to their performance. Only encouragement including statements such as; "Your doing fine" and "Keep going, you're almost finished," were given. Second, I did not deviate from the testing pattern of four trials at each randomly selected wavelength. Such simple deviations as repeating a trial were interpreted by the subjects as an inadequate response on their part and consequently had an impact on the testing session. Third, to reduce experimenter bias I did not edit the data set. All subjects completed one and only one experimental session for each condition and no trials were dropped nor were

any trials repeated or substituted. All calculations were conducted using this unedited data set.

Another concern is the influence of adaptation on visual performance. Adaptation has been an ubiquitous confounding variable since the first experiments on luminous efficiency.

> Adaptation and fatigue are probably the physiological factors most difficult to estimate and control. Color vision is peculiarly dependent on adaptation. When physiologists work on problems of vision they distinguish between the light-adapted and dark-adapted eye depending on whether the eye works in the light or comparative darkness. The ability to perceive color, which is lost on greatly decreasing the illumination, gradually returns as the eye becomes accustomed to the small quantity of light. Fatiguing the eye for one color makes it more sensitive to others. Fatigue also alters the relative critical frequencies of flicker for different colours; and the effect is different depending on the character of the fatiguing light. Both these disturbing factors must be kept in mind when investigating heterochromatic photometry (Ives, 1912).

The broadly sloping spectral curves for the adapting sources are not of particular concern for extrapolation to normal viewing conditions since there are no naturally occurring spectrally flat scenes. Here it is important to keep in mind the premise of photometry, that luminous efficiency functions are moderately robust across conditions. If luminous efficiency functions fluctuated widely with small changes in viewing conditions, photometry would have little practical meaning. Alternatively, if $V(\lambda)$ held strictly across all viewing conditions and adaptation states then there would be no purpose in pursuing luminous efficiency research.

Still, the variability in spectral content with reference intensity does complicate the interpretation of the increase in relative sensitivity at higher intensities by the heterochromatic method. As shown in Figure 3 the short wavelength content of the adapting source increases with increasing reference intensity. Theoretically, the increasing short wavelength content in the adapting source could suppress the short and middle wavelength cone outputs and thus result in an increase in relative sensitivity at longer wavelengths. However there are several pieces of evidence to suggest that the adaptation source is not the cause of the long wavelength (red) shift with increasing intensity in the heterochromatic brightness matching curves.

Visual adaptation to light is generally discussed in terms photopigment bleaching and regeneration at the photoreceptor. The consistency of the peaks in the flicker data at the different intensities argues against adaptation or saturation at photoreceptor as a viable explanation of our heterochromatic brightness matching results. The flicker data do not rule out adaptation or saturation in some location in the visual pathways that is relatively isolated from luminance (flicker) processing.

Relative suppression or saturation would only be expected when the short wavelength output of the adapting source exceeds the long wavelength output. The spectral output curves for the adapting source show that at the 1fl and 10fl levels the long wavelength content is greater than the short wavelength. The 100fl adapting source has a very flat spectral output across the visual spectrum. Only at the 1000fl reference intensity does the short wavelength content exceed the long wavelength. The red shift in the heterochromatic matching curves starts at the 10fl, a level at which there is still a preponderance of long wavelength content. The large peak in the 600nm region is fully

developed at 100fl when the adapting source was most neutral. In addition, in previous work in which adaptive fields have been used to suppress relative sensitivity in different portions of the visual spectrum, Sperling and Harwerth (1971) found that a 5500K correlated color temperature source, very similar the 1000fl adaptive source, to be neutral in terms of relative suppression.

Another issue important to interpretation of these data is inter-subject variability. From the beginning of luminous efficiency research, experimenters have had difficulty with inter-subject variability (Ives, 1912; Gibson and Tyndall 1923). A variety of factors, including differences in optical properties including pupil size, macular pigment density, cone photopigment absorption, neural processing and match criteria are responsible for this inter-subject variability. Experimenters have to decide if they want to control, eliminate or allow these factors to vary. To control for the variability due to pupil size, vision scientists developed a unit of retinal illuminance known as the troland. The troland is not a direct measure of retinal illuminance but is calculated from the source luminance and the subject's pupil size. Efforts to fix pupil size to support a retinal illuminance calculation for a given experiment typically require pupil dilation and/or exact head positioning of subject. Both of these requirements are difficult for subjects in multi-hour experiments and therefore a decision was made to allow pupil diameter to vary in this study. Allowing pupil fluctuation not only improves subject comfort but may also help in the generalization of the experimental findings to real world conditions.

Pupil fluctuation does prevent exact comparison between the present results and those in other studies that have reported in units of retinal illuminance. Pupil fluctuation is also a potential explanation for why some of the subjects start showing greater

sensitivity at longer wavelengths in the heterochromatic brightness condition at lower stimulus levels then others. For example, at the 10 fl intensity subject #4's brightness matching curve has already shown at dramatic shift while at that same level subject #2's brightness matching curve is still very similar to the 1fl curve (Figures 9B and 11B).

Pupil fluctuation may also have a small impact on the spectral content of light reaching the retina and, thus, could explain the narrowing of the flicker photometry curve at the 100fl intensity. The interaction between chromatic aberration and aperture size was one cause for the difficulty in maintaining constant spectral content in this experiment's adaptation stimulus. A similar interaction between the eye's chromatic aberration and its variable aperture, the pupil, may account for our flicker data although Sagawa et al., (1991) had similar flicker photometry findings while using a Maxwellian view optical system that controls for pupil size. The flicker photometry results Sagawa et al. differed from the present study in that they found a small decrease in sensitivity only at longer wavelengths with increasing retinal illuminance. In the present study, there was a decrease in sensitivity at both long and short wavelengths.

Of the studies that can be compared to the current experiment Sagawa et al. (1991) is the most recent and directly relevant. Sagawa et al. conducted both flicker photometry and heterochromatic brightness matching experiments across 3 logs unit (7 levels) of reference intensity. Their brightest level was slightly more intense than the highest level in this study.

In their heterochromatic brightness matching experiments Sagawa and colleagues found that the relative sensitivity of both the longer and shorter wavelengths increased with increasing reference intensity between 100 trolands (~1fl) and 3000 trolands (~50fl).

Sagawa et al. considered the increasing relative sensitivity in the short and long wavelengths to be evidence of increasing chromatic contributions to brightness perception. Between 3000 and 100,000 trolands they found that the relative sensitivities were stable across wavelengths. This finding was interpreted as a saturation of the chromatic channel. In contrast to the results of Sagawa et al., visual inspection of the heterochromatic brightness matching curves in the current study suggests that the relative sensitivity continues to change at higher intensities. The 1000fl curve appears broader and long wavelength relative sensitivity improved, particularly at 650nm, when compared to the 100fl curve. Admittedly this change is small compared to those that occurred in the range over which Sagawa et al. thought that luminous efficiency was actively changing.

At all reference intensities, the heterochromatic brightness matching curves of Sagawa et al. have a definitive unimodal peak near 540nm with a small plateau between 580 and 620nm and have a very similar appearance to the CIE Heterochromatic Luminous Efficiency Curve (CIE HBM Standard) shown in Figure 17. Their data are similar to that from the current study in that for the longer wavelengths the sensitivities from heterochromatic brightness matching increase with increasing reference intensity at least up to 3000 trolands. The magnitude of the increase was less in the subjects studied by Sagawa et al. than in the current study and Sagawa et al. did not find a "notch" between 550 and 600nm as was found in this study. Additionally, in the study of Sagawa et al. the peaks of the heterochromatic brightness matching curves remained at

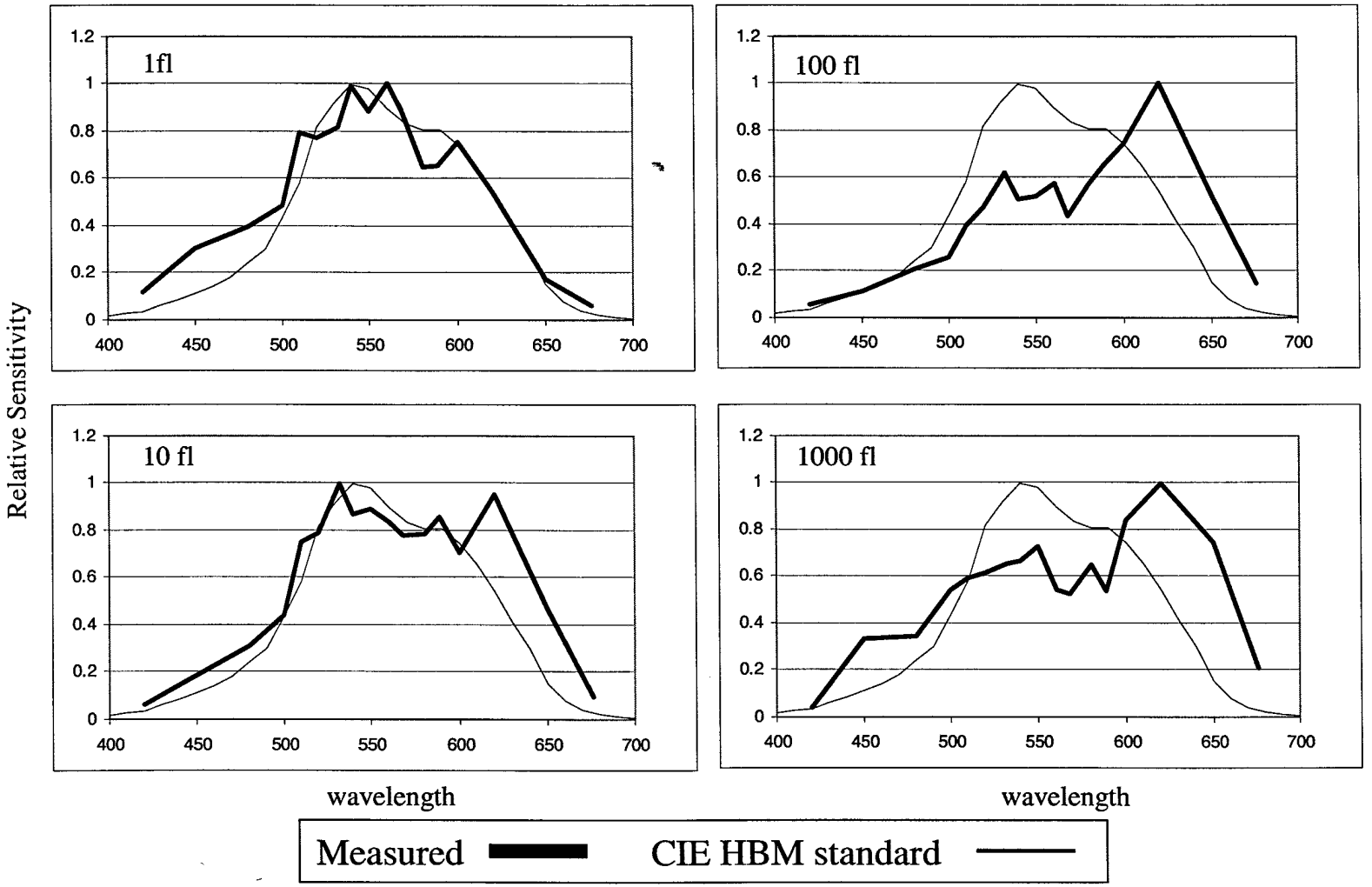

Figure 17: Compares the relative sensitivity measured by heterochromatic brightness matching (Measured) to the CIE luminous efficiency function by heterochromatic brightness matching (CIE HBM Standard) at the different reference intensities. The 1fl curve is consistent with the CIE standard but the higher intensity curves deviate dramatically from the standard

approximately 540nm at all levels of reference intensity. In the present data, the peak sensitivities shifted into the low 600nm region as reference intensity increased.

There are several advantages to the strategy, used in both the study of Sagawa et al. and the present study, of conducting flicker photometry and heterochromatic brightness matching experiments together. The similarity of the flicker photometry curves to $V(\lambda)$ and the relative stability of the curves across the different reference intensities provides reassurance that our apparatus, subjects, calibration processes and testing procedures were appropriate. In addition, both studies were able to compare the sensitivity of the subjects using the two methods at each wavelength. When Sagawa's group plotted the ratio of the heterochromatic brightness matching sensitivity verses the flicker photometry sensitivity, they found that the ratio was greater than one for all wavelengths and that the ratio increased with increasing reference intensity. In the present study a repeated measures t-test was conducted on the flicker photometry and heterochromatic brightness matching data and found that the heterochromatic values were higher than the flicker values at the10fl, 100 and 1000fl levels. This comparison supports the conclusion that the chromatic channel, which is considered an opponent or subtractive process, must produce a positive signal to add to the achromatic channel signal during brightness processing.

Other studies have also shown comparable but less dramatic changes in heterochromatic brightness matching curves relative to the present project. Yaguchi and Ikeda (1980) conducted heterochromatic a brightness matching experiment over 3 log units (4 levels) of reference intensity with their brightest level being approximately equal to our 10fl condition. At their highest intensity level, they concluded that 1 of 4 subjects

showed a dramatic change in relative sensitivity, most notably the development of a second peak at 600nm. One subject showed no change at all, and 2 subjects showed intermediate changes. Overall, they concluded that:

> There are two extreme types among observers at high illuminance levels. One type shows a change in the relative luminous efficiency curve with the change of the retinal illuminance level, and another type shows no change.

In my study, none of the subjects were of the type that showed no change. All 7 subjects showed a change in the luminous efficiency function by the 100fl level, a slightly higher intensity than used by Yaguchi and Ikeda (Figures 8-14).

The major difference between Yaguchi and Ikeda's study and mine is that one of their subjects showed no change in heterochromatic brightness matching curves with increasing reference intensity, while I found none. There are two potential explanations for this difference. It is possible that there is a population of individuals whose brightness matching curves are stable across all photopic reference intensities and that my subject pool failed to include any of these "no change" type of subjects. However, the more likely explanation is that most subjects show an increase in relative sensitivity at longer wavelengths provided the reference level is sufficiently high.

Insufficient reference intensity is also the likely explanation for the differences between the heterochromatic brightness matching curves in the present study and the Luminous Efficiency Curve for Centrally-Viewed, Two Degree Field by Heterochromatic Brightness Matching Function published by CIE (Meyer et al, 1978). The function is an average taken from 31 subjects across 7 different studies (Bedford and Wyszecki, 1958;

Comerford and Kaiser, 1975; Guth and Lodge, 1973; Kinney 1964; Sperling and Lewis, 1959; Wagner and Boynton, 1972). Reference intensities for these studies were reported to be 500 trolands (~10fl) or less except for one that seemed to be conducted at a fairly bright level but for which no reference intensity was reported.

The CIE heterochromatic brightness luminous efficiency curve is shown with each of my measured heterochromatic brightness matching curves in Figure 17. At 1fl, my measured heterochromatic brightness matching curve agrees well with the CIE heterochromatic brightness curve. However, by the 10fl level, differences between the CIE standard and my measurements become apparent in the long wavelength portion of the spectrum. At 100fl and 1000fl, the present measurements are clearly disparate from the CIE "standard." This disparity suggests that the CIE standard may seriously underestimate the relative sensitivity of the eye in the long wavelength region of the visible spectrum at higher intensities.

A bipartite color match is very similar to heterochromatic brightness matching except that three color primaries are used to match the reference stimulus and the match is made on both hue and brightness. The laws of brightness matching imply that, if a color match is made between 3 monochromatic primaries and a reference stimulus of fixed relative spectral content, the relative proportions of the primaries will remain the same regardless of the luminance of the reference stimulus. Wyszecki and Stiles (1980) found that proportionality failed for trichromatic matches when the reference intensities were in the high photopic range. They used a pigment bleaching hypothesis to model their results and found reasonable agreement between the model and their data. There are two interesting comparisons between their study and my results. The first involves the

intensity levels at which the non-linearities occurred. Wyszecki and Stiles found the loss of proportionality occurred between 600 and 5000 fl for their subject. For at least three of my subjects, the heterochromatic brightness matching curves had started to change shape at the 10fl level. From this it would seem that the violations in the laws of brightness matching started at least one order of magnitude lower for the present experiment. The second comparison concerns pigment bleaching. Wyszecki and Stiles felt that pigment bleaching was a good explanation for the failure of proportionality. However, in the present study the constancy in the flicker photometry curves across the reference intensities argues against a photoreceptor explanation for the violation of the brightness matching laws under the heterochromatic brightness matching condition.

A better explanation for the finding from this current study may be found in the work of Sperling and Harwerth (1971). They conducted relative sensitivity studies over 3 log units of adaptation intensity using an increment-threshold approach. They concluded that the increment threshold relative sensitivity curves had three peaks (at 445nm, 540nm and 610nm) and that the two longer wavelength peaks were best described by a linear subtractive model, where the middle and long wavelength photoreceptors have mutually inhibitory relationship. There are substantial similarities between the data obtained in this study and those of Sperling and Harwerth. In the present study, every subject has a peak in their heterochromatic brightness matching curves at or above 600nm for the 100fl and 1000fl intensities (Figures 8-14). These peaks are large enough that on the mean curves for heterochromatic brightness matching at 100fl and 1000fl (Figure 15), the absolute maximum is above 600nm. The exact location of the peak cannot be determined because of the spacing between the filters used for this region of the spectrum. Figure 15

also shows peaks near 540nm for all intensities, but little evidence of peaks near 445nm as found by Sperling and Harwerth.

An important finding from the Sperling and Harwerth study was that the amount of interaction or inhibition between the middle and long wavelength cone responses increased with increasing background luminance. This increased interaction is shown in the relative sensitivity curve in a deepening notch at 580nm as the background luminance is increased. With a 10 troland (~1fl) background, their sensitivity curves had a broad single-peaked appearance with a slight depression in the 580nm region. This description is also reasonably representative of the mean heterochromatic brightness matching curve generated at a similar intensity in the present experiment (1fl, Figure 15). When the background was 1000 trolands, roughly equivalent to my 100fl level, Sperling and Harwerth found a dramatic three-peaked appearance with a deep notch at 580nm. At 10000 trolands, the three peaked shape was even more pronounced with the middle and long wavelength peaks being of comparable size. As described in the results section, the mean curves from the current study produce similar middle and long wavelength peaks with a notch around 580nm. However, I did not find evidence of the short wavelength peak. The striking similarity, thus, is that, in both studies, the sensitivity curves for heterochromatic brightness matching go through a similar transformation in the middle and long wavelength regions at similar levels of reference intensity.

While my heterochromatic brightness matching results are similar to those of several previous studies, one finding in the flicker data is distinct from previous work: the significant difference between the 100fl flicker curve and the other levels. It is generally assumed that data obtained by flicker photometry reflect processing by an achromatic or

luminance channel within the visual system. The luminance channel is assumed to be very predictable, supporting the brightness matching laws of symmetry, transitivity, proportionality and additivity. It is also assumed that the luminance channel is derived from straightforward processing of the output of retinal cones. The significant probability of Level for the flicker data suggests that these assumptions do not strictly hold. Two notable qualifications should be made about this conclusion. First, in the $V(\lambda)$ style transformation of the data, the peak values for each curve are given more weight than the others. This is an infringement on the assumptions underlying analysis of variance and therefore the resulting probabilities should be considered with some degree of suspicion. Second, the size of the effect is small and thus is not likely to have substantial practical influence on brightness processing.

Even if these reservations are acknowledged, highlighting the possible failure of brightness matching laws at 100fl is important. In science the most parsimonious theory that accurately explains the data is preferred. If flicker curves are consistent across intensities, a relatively simple luminance channel model, using the summation of middle and long wavelength cone outputs, can be used to fit the data. However, if the relative sensitivity by flicker photometry changes with intensity, a more complicated model is needed. Research showing that short wavelength cones can influence luminance channel processing at higher adaptation levels may provide a potential foundation for future luminance models.

Stockman et al (1990) conducted flicker photometry trials using a variety of monochromatic stimuli and backgrounds to isolate short, middle and long wavelength cone pathways. Using this approach they were able to make short wavelength stimuli

flicker matches to middle and long wavelength stimuli for frequencies up to 30 Hz, thereby demonstrating a short wavelength cone contribution to luminance processing. In addition, they found that to optimize the short wavelength contribution to the flicker match, a phase lag had to be introduced, thus demonstrating that the processing time lag for flicker is different for each cone pathway. By investigating the relationship between flicker rates and phase lags Stockman et al were also able to demonstrate that the usual blue cone luminance signal is antagonistic and that the signal can interact differently with the middle and long wavelength signals depending on flicker rates and phase lag. Since stimulus intensity is known to influence processing time it is reasonable to believe that the changes in reference intensity in the current study could influence the interactions between the three cone pathways in luminance processing.

The results of the present study may thus highlight potential neural processing anomalies. The more immediate and practical outcomes of this study are with regard to photometry at higher intensities. 1912, Herbert Ives proclaimed: "Given the satisfactory photometric method, and the standard eye to use it, the measurement of the relative brightness of different coloured light becomes a definite thing." He had just completed luminous efficiency curves for 18 subjects and felt that the mean for these subjects was "sufficiently near the mean eye for all practical purposes." Ives was confident he had resolved the photometry issue: "The results of the photometric method here advocated hold for the average eye under the most important illumination conditions." However, Ives acknowledged that some adjustments to the photometric system would need to be made for unusual illuminations. "It is inevitable that correction factors will need to be

applied to these values whenever the effective illuminations are widely different from that here adopted as standard" (Ives, 1912).

It is evident in my study that the effective illuminations created by simultaneously bright and narrow light sources such as laser and LEDs can differ greatly from those used to develop V(λ). Though less well known than V(λ), the CIE Luminous Efficiency Curve for Centrally-Viewed, Two Degree Field by Heterochromatic Brightness Matching is the preferred function for photometry of narrow sources (Meyer et al, 1978). The present study agrees well with the CIE heterochromatic brightness matching function at the 1fl reference intensity. However at 10 foot-lamberts or greater, the present study suggests that at luminance of the CIE curve may substantially under estimate the luminous efficiency of long wavelengths.

My results create some ambiguity as to how LED and laser photometric measurements should be made at higher intensities. This ambiguity will continue until further experimentation can place these results in context. A practical solution, in the mean time, would be to consider the luminous efficiency curve at high intensities to be flat over most of the spectrum, from about 480nm to 620nm, and to acknowledge that subject variability makes prediction of individual performance problematic. More generally, it is important to acknowledge the quality of our measurements, an issue that is often overlooked. As pointed out by W.D. Wright who was instrumental in establishing the 1931 CIE system:

> The CIE Colorimetry Committee recently in their wisdom have been looking at the old 1931 observer and have been smoothing the data to obtain more consistent calculations with computers. This has also involved some extrapolation and, in smoothing, they have added some

additional decimal places. When I look at the revised table of the x,y,z, functions, I am rather surprised to say the least. You see, I know how accurate the actual measurements really were. Guild did not take any observations below 400nm and neither did I, and neither did Gibson and Tyndall on the V(λ) curve, and yet at a wavelength of 362nm, for example, we find a value of y of .000004929604! This, in spite of the fact that at 400nm the value of y may be in error by a factor of 10. I can not help wondering what Mr. Guild thinks if he happens to see these tables. I know we can put the blame on the computer but we must not abdicate our common sense altogether (Boyton, 1979).

Conclusion

In this study luminous efficiency curves were generated using two methods, flicker photometry and heterochromatic brightness matching, at four levels of reference intensity, 1, 10, 100, and 1000 fl. Luminous efficiency was found to vary with both the method and intensity with which the measurement was taken. At the 1fl reference intensity, the heterochromatic brightness matching and flicker photometry curves were similar to each other and to V(λ). As reference intensity was increased, the heterochromatic brightness matching relative sensitivity in the long wavelength region of the visible spectrum increased and the peak luminous efficiency shifted from approximately 540nm to above 600nm.

The most reasonable explanation for the long wavelength shift is that, as the intensity increases, the contribution of the opponenency based chromatic system increases. Similar results have been reported in other heterochromatic brightness matching studies and by Sperling and Harwerth using the increment threshold technique. "Thus as the neutral adaptive state of the fovea is increased, the amount of interaction between the red- and green-receptor response increases" (Sperling and Harwerth 1971).

The change in the flicker photometry curve was less dramatic. The peak in the flicker photometry luminous efficiency curve remained at approximately 550nm at all reference intensities but the luminous efficiency in the both the long and short wavelength regions appeared to drop slightly at the 100fl reference intensity.

The variability in luminous efficiency functions across levels of photopic adaptation has implications for both the study of visual processing and the photometric measurement of simultaneously bright and spectrally narrow light sources that are increasingly prevalent in our environment.

References

Alpern, M., Rushton, W., Torii, S. Signals from Cones. Journal of Physiology. 207, 463-475 (1970)

Bedford, R., Wyszecki, G. Luminosity Functions for Various Field Sizes and Levels of Retinal Illuminance. Journal of the Optical Society of America. 48, 406-411 (1958)

Boyton, R. Human Color Vision. Holt, Rinehart and Winston (1979)

Comerford, J., Kaiser, P. Luminous Efficiency Functions by Heterochromatic Brightness Matching. Journal of the Optical Society of America. 65, 466-468 (1975)

De Vries, H. The Luminosity Curve of the Eye as Determined by Measurements with the Flickerphotometer. Physica. 14, 319-348 (1948)

Gibson, K.,Tyndall, E. Visibility of Radiant Energy. Scientific Papers of the Bureau of Standards No. 475, Department of Commerce (1923)

Guth, S., Lodge, H. Heterochromatic Additivity, Foveal Spectral Sensitivity, and a New Color Model. Journal of the Optical Society of America. 63, 450-462 (1973)

Ikeda, M., Shimozono, H. Luminous Efficiency Functions Determined by Successive Brightness Matching. Journal of the Optical Society of America. 68, 1767-1771 (1978)

Ikeda, M., Yaguchi, H., Sagawa, K. Brightness Luminous-Efficiency Functions for 2 degree and 10 degree Fields. Journal of the Optical Society of America. 72, 1660-1665 (1982)

Ives, H. Studies in the Photometry of Lights of Different Colours. I. Spectral Luminosity Curves Obtained by the Equality of the Brightness Photometer and the Flicker Photometer Under Similar Conditions. Philosophical Magazine 24, 149-188 (1912)

Ives, H. Studies in the Photometry of Lights of Different Colours. V. The Spectral Luminosity Curve of the Average Eye. Philosophical Magazine 24, 853-863 (1912)

Johnston, S. A History of Light and Color Measurement. Institute of Physics Publishing. 2001

Kaiser, P. CIE 1988 2 ° Spectral Luminous Efficiency Function for Photopic Vision. Publication CIE 86-1990 (1990)

Kinney J. Effect of Field Size and Position on Mesopic Spectral Sensitivity. Journal of the Optical Society of America. 54, 671-677 (1964)

Kinney, J. Brightness of Colored Self-Luminous Displays. Color Research and Application. 8, 82-89 (1983)

King-Smith, P., Carden, D. Luminance and Opponent-Color Contributions to Visual Detection and Adaptation and Temporal and Spatial Integration. Journal of the Optical Society of America. 66, 709-717 (1976)

Lennie, P., Pokorny, J., Smith, V. Luminance. Journal Of The Optical Society Of America A-Optics Image Science And Vision. 10, 1283-1293 (1993)

Meyer, E. et al. Light as a True Visual Quantity: Principles of Measurement. Publication CIE 41 (TC-1.4). International Commission of Illumination (1978)

Pridmore, R. Bezold-Brucke Hue-Shifts as Functions of Luminance Level, Luminance Ratio, Interstimulus Interval and Adapting White for Aperture and Object Colors. Vision Research. 39, 3873-3891 (1999)

Science and Engineering Associates Inc http://www.seabase.com/page.asp?PageId=82 April 30, 2002

Sagawa, K., Yamada, M., Yaguchi, H., Miyake, Y. Luminance Efficiency Functions at High Luminance Levels. CIE Proceedings, 22nd Session-Division 1, 69-70 (1991)

Sliney, D., Wolbarsht, M. Safety with Lasers and Other Optical Sources. Plenum Press (1980)

Southall, P. Helmholtz's Treatise on Physiological Optics. Dover Publications (1962)
Sperling H., Lewis, G. Some Comparisons Between Foveal Sensitivity Data Obtained at High Brightness and Absolute Threshold. Journal of the Optical Society of America. 49, 983-989 (1959)

Sperling, H., Harwerth R. Red-Green Cone Interactions in the Increment-Threshold Spectral Sensitivity of Primates. Science. 172, 180-184 (1971)

Stevens, S. Handbook of Experimental Psychology. John Wiley & Sons (1951)

Stockman, A., MacLeod, D., DePriest, D. The Temporal Properties of the Human Short-Wave Photoreceptors and Their Associated Pathways. Vision Research. 31, 189-208 (1991)

Thomson, L. The Effect of Change of Brightness Level upon the Foveal Luminosity Curve Measured with Small Fields. Journal of Physiology. 106, 368-377 (1947)

Vos, J.,Friele, L., Walranven, P. Color Metrics. AIC/Holland, 2-3 (1972)

Wagner, G., Boynton, R., Comparison of Four Methods of Heterochromatic Photometry. Journal of the Optical Society of America. 62, 1508-1515 (1972)

Wyszecki, G., Stiles, W. High-Level Trichromatic Color Matching and the Pigment-Bleaching Hypothesis. Vision Research. 20, 23-37 (1980)

Wyszecki, G., Stiles, W. Color Science Concepts and Methods, Quantitative Data and Formulae, 2nd Edition. John Wiley & Sons (1982)

Yaguchi, H., Ikeda, M. Luminous Efficiency Functions by Heterochromatic Brightness Matching for a Wide Range of Retinal Illuminance. J. Light & Vis. Env. 4, 14-17 (1980)